



```

67 uAspAspLeuValAsnTrpAlaArgSerSerLeuTrpProMetThrP 84
|||||
221 GGTACCTCGTCAACTGGGCGCGGGAGTTCTGTGGCCATGACCT 270
|||||
84 heGlyLeuAlaCysCysAlaValGluMetMetHisMetAlaAlaProArg 100
|||||
271 TCGGCTGGCCTGCTGCGCGGTGGAGATGATGCATGGCAGCAGCCCGC 320
|||||
101 TyrAspMetAspArgPheGlyValValPheArgAlaSerProArgGlnSe 117
|||||
321 TAGGACATGACCGCTTTGGCGTGTCTTCCGCGCAGCGCGCGCATG 370
|||||
117 rAspValMetIleValAlaGlyThrLeuThrAsnLysMetAlaProAlaL 134
|||||
371 CGAGCTCATGATCGTGGCGCGGACACTACCAACAAGATGGCCCGCAGC 420
|||||
134 euArgLysValTyrAspGlnMetProGluProArgTyrValValSerMet 150
|||||
421 TTCGCAAGGTCTACGACCATGTCGGGAGCGCGCTAGTGTGTCTCCATG 470
|||||
151 GlySerCysAlaAsnGlyGlyTyrTyrHisTyrSerTyrSerValVa 167
|||||
471 GAGAGCTGGCGCCCAACGAGGAGGCTACTACCACTATTCTCTCGTGTGT 520
|||||
167 lArgGlyCysAspArgIleValProValAspIleTyrIleProGlyCysP 184
|||||
521 GAGGCGCTGGCAGCCGATCGTGGCGGTGGACATCTACATCCAGGCTGCC 570
|||||
184 roProThrAlaGluAlaLeuLeuTyrGlyIleLeuGlnLeuArgLys 200
|||||
571 CAGCTACGGCGGAGGCGCTGCTCTAGGCGATCTGCAGTGCAGAGGAAG 620
|||||
201 lLysArgGluArgArgLeuGlnIleTyrTyrArgArg 213
|||||
621 ATCAAGCGGAGCGGAGGCTGACGATCTGTGTACCGCAGG 659

```

seq\_name: gb\_pr:AF060512

```

seq_documentation_block:
LOCUS AF060512 754 bp mRNA PRI 02-JAN-2001
DEFINITION Homo sapiens clone 016d03 My017 protein mRNA, complete cds.
ACCESSION AF060512
VERSION AF060512.1 GI:12001973
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (Bases 1 to 754)
AUTHORS Mao,Y.M., Xie,Y., Huang,X.Y., Ying,X. and Dai,J.L.
TITLE Direct Submission
JOURNAL Submitted (20-APR-1998) Institute of Genetics, School of Life
Science, Fudan University, 220 Handan Rd., Shanghai 200433,
P. R. China

```

```

FEATURES
source
1..754
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="016d03"
/tissue_type="brain"
/dev_stage="fetus"
10..654
/codon_start=1
/product="My017 protein"
/protein_id="AAG43130.1"
/db_xref="GI:12001974"
/translation="MAVLSAPGLRFRILGLRSSVLAYOARGVHOSVATDGPSSNOP
ALPARAVAPKPSRGEYVAKLLDLYNWRRLSPFPGIACCAVEMHWAAPYD
MDRGVFRASPROSDVNIIVAGILTNWAPLRKSPRPDAGALRGLHGLRQGRIL
LPFLLGEGRLPHRARGHLRPLTFAEALLYLQILQRLIKRRRLQIWRKR"
BASE COUNT 120 a 267 c 239 g 128 t
ORIGIN

```

```

alignment_scores:
Quality: 1046.50 Length: 215
Ratio: 4.960 Gaps: 2
Percent Similarity: 98.140 Percent Identity: 97.209

alignment_block:
US-09-525-867-1 x AF060512
..
Align seg 1/1 to: AF060512 from: 1 to: 754

1 MetAlaValLeuSerAlaProGlyLeuArgGlyPheArgIleLeuGlyLe 17
|||||
10 ATGGCGGTGCTGTACGCTCTGGCTGCGCGCTTCCGGATCCTTGGTCT 59
|||||
17 uArgSerSerValGlyProAlaValGlnAlaArgGlyValHisGlnServ 34
|||||
60 GCCTCCAGCGTGGCGCTGGCTGTGCAGCAGCAGGCTGTCCATCAGAGC 109
|||||
34 alAlaThrAspGlyProSerSerThrGlnProAlaLeuProLysAlaArg 50
|||||
110 TGCCACCATGATGCCCAAGCAGCAGCCAGCTGCCCTGCCAAGGCCAGA 159
|||||
51 AlaValAlaProLysProSerSerArgGlyGluTyrValValAlaLysLe 67
|||||
160 GCCGTGGCTCCCAACCCAGCAGCGCGGCGAGTATGTGTGGCCAAAGCT 209
|||||
67 uAspAspLeuValAsnTrpAlaArgSerSerLeuTrpProMetThrP 84
|||||
210 GGTACCTCGTCAACTGGGCGCGCGGAGTTCTGTGGCCCATGACCT 259
|||||
84 heGlyLeuAlaCysCysAlaValGluMetMetHisMetAlaAlaProArg 100
|||||
260 TCGGCTGCGCTGCTGGCGGTGGAGATGATGCACATGGCAGCAGCCCGC 309
|||||
101 TyrAspMetAspArgPheGlyValValPheArgAlaSerProArgGlnSe 117
|||||
310 TAGGACATGACCGCTTTGGCGTGTGTTCCGCGCAGCGCGCGCAGTCTC 359
|||||
117 rAspValMetIleValAlaGlyThrLeuThrAsnLysMetAlaProAlaL 134
|||||
360 CGAGCTCATGATCGTGGCGCGCACACTACCAACAAGATGGCCCGCAGCGC 409
|||||
134 euArgLys.Val...TyrAspGlnMetProGluProArgTyrValValSe 149
|||||
410 TTGCAAGTCTCCCCCAGCAGCAGATGCCGAGCGCGCTACGTGGTCTC 459
|||||
149 rMetGlySerCysAlaAsnGlyGlyTyrTyrHisTyrSerTyrServ 166
|||||
460 CATGGGAGCTGCGCCACGCGGAGGCTACTACCACTATTCTTACTCGG 509
|||||
166 alValArgGlyCysAspArgIleValProValAspIleTyrIleProGly 182
|||||
510 TGTGAGGGCTCGACCGCATCTGCTCCCGTGGACATCTACATCCAGGC 559
|||||
183 CysProProThrAlaGluAlaLeuLeuTyrGlyIleLeuGlnLeuGlnAr 199
|||||
560 TGCCCACTACGCCGAGGCGCTGTCTACGCGCATCTCGACATCTACATCCAGAG 608
|||||
199 gLysIleLysArgGluArgGlnLeuGlnIleTyrTyrArgArg 213
|||||
609 GAAGATCAAGCGGAGCGGAGGCTGCAGATCTGTGTACCGCAGG 651

```

seq\_name: gb\_on:BTPSSTSU

```

seq_documentation_block:
LOCUS BTPSSTSU 752 bp mRNA MAM
DEFINITION B.taurus mRNA for PSST subunit of the NADH: ubiquinone
oxidoreductase complex.
ACCESSION X65020
VERSION X65020.1 GI:11256
KEYWORDS NADH-ubiquinone oxidoreductase subunit.
SOURCE cow.

```

29-NOV-1994

```

ORGANISM      Bos taurus
               Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
               Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
               Bovidae; Bovinae; Bos.
REFERENCE      1 (bases 1 to 752)
AUTHORS        Arizumi, J.M., Runswick, M.J., Skehel, J.M. and Walker, J.E.
TITLE          NADH: ubiquinone oxidoreductase from bovine heart mitochondria. A
               fourth nuclear encoded subunit with a homologue encoded in
               chloroplast genomes
JOURNAL        FEBS Lett. 301 (3), 237-242 (1992)
MEDLINE        9249573
REFERENCE      2 (bases 1 to 752)
AUTHORS        Walker, J.E.
TITLE          Direct Submission
JOURNAL        Submitted (11-MAR-1992) J.E. Walker, M.R.C.Lab of Molecular
               Biology, Hills Road, Cambridge CB2 2QH, UK
FEATURES       source
               1..752
               /organism="Bos taurus"
               /db_xref="taxon:9913"
               /germline
               /rev_stage="adult"
               /tissue_type="heart"
               17..667
               /gene="ci-PSST"
               17..667
               /partial
               /gene="ci-PSST"
               /EC_number="1.6.99.3"
               /codon_start=1
               /evidence=experimental
               /product="PST subunit of the NADH: ubiquinone
               oxidoreductase complex"
               /protein_id="CAA46154.1"
               /db_xref="GI:599691"
               /db_xref="SWISS-PROT:P42026"
               /translation="WAALAAELRLPILAVSGVGAQLQVRCVHSSMAADSPSTQPA
               VSGRAVVPKPAALPSSRGVEYVAKLDDLINWARRSLWMPFGLACCAVEMHMAAP
               RYDMDFGVFRASPDQDMIVAGLTNKPALRKVYDQMPPEPRYVWNGSCANG
               GYHYSYVVRGCDRIVPDYIVPGCPPTAEALLYILQLOKIKRKLRIWYR"
               128..664
               /gene="ci-PSST"
               /product="PST subunit of the NADH: ubiquinone
               oxidoreductase complex"
               714..719
               /evidence=experimental
BASE COUNT    145 a 243 c 230 g 134 t
ORIGIN
alignment_scores:
  Quality: 965.50      Length: 217
  Ratio: 4.710         Gaps: 2
  Percent Similarity: 94.470 Percent Identity: 85.714
alignment_block:
US-09-525-867-1 x BTPSSTU ..
Align seg 1/1 to: BTPSSTU from: 1 to: 752
1 MetAlaValLeuSerAlaProGlyLeuArgGlyPheArgIleLeuGlyLe 17
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
17 ATGGGGGGCGTGGCGCTCTCCGCTGCTC...CACCTATCTTCTGT 63
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
17 uArgSerValGlyProAlaValGlnAlaArgGlyValHisGlnSerV 34
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
64 GGGCTCTGGTGGTGGCGCTCTGCAGTGCAGGGGTCATCTAGCA 113
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
34 aAlaThrAspGlyProSerSerThrGlnProAlaLeuProLysAlaArg 50
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
114 TGGCCGCTGACAGCCGAGCAGCACTAGCCCGCGGTATCCAGGGCAGA 163

```

```

51 AlaValAlaProLys.....ProSerSerArgGlyGluTyrVa 63
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
164 GCTGTGTCCCAACCTCCGCGCACTTCCAGCAGCGCGGCGAGTATGT 213
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
63 lValAlaLysLeuAspAspLeuValAsnTrpAlaArgArgSerSerLeu 80
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
214 GGTGCCAAGCTGGACACCTCATCACTGGCGCGCGGAGCTCGCTGT 263
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
80 rpProMetThrPheGlyLeuAlaCysCysAlaValGluMetMetHisMet 96
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
264 GGCACATGACCTTTGGCTGGCTGCTGCTGCTGGAGATGATGCACATG 313
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
97 AlaAlaProArgTyrAspMetAspArgPheGlyValValPheArgAlase 113
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
314 CGCGCACCCGCTATGACATGGACCGCTTTGGCGTGGTCTTCCGTGCCAG 363
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
113 rProArgGlnSerAspValMetIleValAlaGlyThrLeuThrAsnLysM 130
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
364 CCCGGCCAGTCCGACGTGATGTTGTGGTGGAGCGCTCACCACAAGA 413
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
130 etAlaProAlaLeuArgLysValTyrAspGlnMetProGluProArgTyr 146
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
414 TGGCCCTTGCCTCCGCAAGTCTACGACCATGCGGAGCCCGCTAT 463
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
147 ValValSerMetGlySerCysAlaAsnGlyGlyTyrTyrHisTyrSe 163
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
464 GTCGTATCCATGGGAGCTGTGCCAACGGAGTGGCTACTACCACACTC 513
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
163 rTyrSerValValArgGlyCysAspArgIleValProValAspIleTyrI 180
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
514 CTACTCAGTGGTGGGCGTGGCGACCGCATCGTTCCAGTGGAGCATCTACG 563
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
180 leProGlyCysProProThrAlaGluAlaLeuLeuTyrGlyIleLeuGln 196
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
564 TCCGAGGCTGCCCGCTACGGCCGAGGCGCTGCTGTATGATTCATTCGAG 613
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
197 LeuGlnArgLysIleLysArgGluArgArgLeuGlnIleTrpTyrArgAr 213
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
614 CTGCAGAAGAAGATCAACCGGAAAGAGGTTCGGATCGTGTACCCGAG 663
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
213 g 213
564 G 664
seq_name: gb.htg:AC014169
seq_documentation_block:
LOCUS      AC014169      29563 bp      DNA      HTG      16-NOV-1999
DEFINITION Drosophila melanogaster, *** SEQUENCING IN PROGRESS ***, in ordered
            pieces.
ACCESSION  AC014169
VERSION    AC014169.1  GI:6437166
KEYWORDS   HTG; HTGS_PHASE2.
SOURCE     fruit fly.
ORGANISM   Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
            Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
            Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE  1 (bases 1 to 29563)
AUTHORS    Adams, M. and Venter, J.C.
TITLE      Direct Submission
JOURNAL    Submitted (16-NOV-1999) Celera Genomics, 45 West Gude Drive,
            Rockville, MD, USA
COMMENT    This sequence was identified as CDM:10210747 by the submitter.
            For further information on this sequence e-mail to fly@celera.com.
            * NOTE: This is a 'working draft' sequence.
            * by the finished sequence as soon as it is available and
            * the accession number will be preserved.
FEATURES   Location/Qualifiers
            source
            1..29563
            /organism="Drosophila melanogaster"
            /db_xref="taxon:7227"

```

BASE COUNT 8656 a 6105 c 6205 g 8597 t  
ORIGIN

## alignment\_scores:

Quality: 756.50 Length: 197  
Ratio: 4.274 Gaps: 3  
Percent Similarity: 89.848 Percent Identity: 71.574

## alignment\_block:

US-09-525-867-1 x AC014169 ..

Align seg 1/1 to: AC014169 from: 1 to: 29563

18 ArgSerValGlyProAlaValGlnAlaArgGlyValHisGlnSerVa 34  
19629 CGCTCTGAATTTGGCCCTGGT.....CGGCCAACAAACGGTTCAG 19672  
34 lAlaThrAspGlyProSerSerThrGlnProAlaLeuProLysAlaArg 51  
19673 TTGCCGAGTGGCCCAAAATCTGCCCAAGAGGCTA.....CTCT 19713  
51 lavalAlaProLysProSerSerArgGlyGluTyrValValAlaLysLe 67  
19714 CGGTTCGGCACCAACAGTCTCCCTTCGGAGTGGTCACTGGCCAGACT 19763

67 uAspAspLeuValAsnTtpAlaArgArgSerLeuTyrProMetThrp 84  
19764 GACGATGCTGCTCACTGGGTTCGCAAGGTCGATCTGGCCACTGACT 19813  
84 heGlyLeuAlaCysCysAlaValGluMetMetHisMetalAlaProArg 100  
19814 TCGTTTGGCTGCTGCTGCGTCGCAATGATGCACATCGTGTCCGCGT 19863  
101 TyrAspMetAspArgPheGlyValValPheArgAlaSerProArgGlnSe 117  
19864 TAGCAGATGATCGATAGTGTGTGTTCGTGCTGCTCCACGTCAAGC 19913  
117 rAspValMetIleValAlaGlyThrLeuThrAsnLysMetalProAlaL 134  
19914 CGATGTCATCATCGTCTGTCGACGCTGACCAACAAATGACCCGCGCC 19963

134 euArgLysValTyrAspGlnMetProGluProArgTyrValValSerMet 150  
19964 TCGGAAGTCTTACGACCAATATGCCGACGACCGTGGGTCTATCTCCATG 20013  
151 GlySerCysAlaAsnGlyGlyTyrTyrHisTyrSerTyrSerValVa 167  
20014 GGCAGTGTGCCAAGCGGGCGGCTACTACCATCTACTCTCGTCTG 20063  
167 lArgGlyCysAspArgIleValProValAspIleTyrIleProGlyCysP 184  
20064 CGGTGGCTCGCATAGATAATCCCGTCGACATATACGTACCCGCTGTC 20113  
184 roProThrAlaGluAlaLeuLeuTyrGlyIleLeuGlnLeuGlnArgLys 200  
20114 CGCCACCCCGGAGGCGCTCATGTACGGCGTTTGCAGCTCAGAGAAG 20163  
201 lIeLysArgGluArgLeuGlnIleTyrTyrArgArg 213  
20164 GTTAAGCGATGAAGACGCTCCAGATGTGGTATAGGAAG 20202

seq\_name: gb\_htg:AC010919

## seq\_documentation\_block:

LOCUS AC010919 140685 bp DNA 13-DEC-1999  
DEFINITION Drosophila melanogaster chromosome x clone BACR32K23 (D891) RPCR-98  
32.K.23 map 14A-14A strain y; on bw sp. \*\*\* SEQUENCING IN PROGRESS  
\*\*\*, 77 unordered pieces.  
ACCESSION AC010919  
VERSION AC010919.9 GI:6563419  
KEYWORDS HTG; HTGS-PHASE1.  
SOURCE fruit fly.

## ORGANISM

Drosophila melanogaster

Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;  
Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.

## REFERENCE

## AUTHORS

1 (bases 1 to 140685)

Celniker,S.E., Agbayani,A., Arcaina,T.T., Baxter,E., Blazej,R.G.,  
Butenhoff,C., Champe,M., Chavez,C., Chew,M., Ciesiolka,L.,  
Doyle,C.M., Farfan,D.E., Galle,R., George,R.A., Harris,N.L.,  
Hinkle,A., Hoskins,R.A., Houston,K.A., Hummasti,S.R., Karra,K.,  
Kearney,L., Lee,B., Lewis,S., Li,P., Ling,H., Moshrefi,A.R.,  
Moshrefi,M., Nixon,K., Pacleb,J.M., Park,S., Pfeiffer,B.,  
Richards,S., Sethi,H., Svirskas,R.R., Wan,K.H., Webster,D.,  
Woolley,P., Yang,S., Yee,M., Yu,C. and Rubin,G.M.

Sequencing of Drosophila melanogaster

## TITLE

## JOURNAL

## REFERENCE

## AUTHORS

Unpublished

2 (bases 1 to 140685)

Celniker,S.E., Agbayani,A., Arcaina,T.T., Baxter,E., Blazej,R.G.,  
Butenhoff,C., Champe,M., Chavez,C., Chew,M., Ciesiolka,L.,  
Doyle,C.M., Farfan,D.E., Galle,R., George,R.A., Harris,N.L.,  
Hoskins,R.A., Houston,K.A., Hummasti,S.R., Karra,K., Kearney,L.,  
Kim,E., Lee,B., Lewis,S., Li,P., Lomotan,M.A., Mazda,P.,  
Moshrefi,A.R., Moshrefi,M., Nixon,K., Pacleb,J.M., Park,S.,  
Pfeiffer,B., Poon,B., Sequeirra,A., Sethi,H., Snir,E.,  
Svirskas,R.R., Wan,K.H., Weinburg,T., Zhang,R., Zieran,L.I. and  
Rubin,G.M.

## Direct Submission

## TITLE

## JOURNAL

## COMMENT

Submitted (27-SEP-1999) Drosophila Genome Center, Lawrence Berkeley  
Laboratory, MS 64-121, Berkeley, CA 94720, USA

On Dec 13, 1999 this sequence version replaced gi:6449494.  
For further information about this sequence, including its location  
and relationship to other sequences, please visit our sequence  
archive Web site (<http://www.fruitfly.org/sequence/>) or send email  
to [bdgp@fruitfly.berkeley.edu](mailto:bdgp@fruitfly.berkeley.edu). All contigs in this submission meet  
the following cutoffs: length >= 200 bases.

\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 77 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.

1 674: contig of 674 bp in length  
\* 675  
\* 754: gap of unknown length  
\* 755  
1507: contig of 753 bp in length  
\* 1508  
1587: gap of unknown length  
\* 1588  
2889: contig of 1302 bp in length  
\* 2890  
2969: gap of unknown length  
\* 2970  
3554: contig of 585 bp in length  
\* 3555  
3634: gap of unknown length  
\* 3635  
4239: contig of 605 bp in length  
\* 4240  
4319: gap of unknown length  
\* 4320  
4698: contig of 379 bp in length  
\* 4699  
4778: gap of unknown length  
\* 4779  
5295: contig of 517 bp in length  
\* 5296  
5375: gap of unknown length  
\* 5376  
5935: contig of 560 bp in length  
\* 6015  
6560: contig of 545 bp in length  
\* 6561  
6640: gap of unknown length  
\* 7460  
7460: gap of unknown length  
\* 7461  
9170: contig of 1630 bp in length  
\* 9171  
9250: gap of unknown length  
\* 9251  
9882: contig of 632 bp in length  
\* 9883  
9962: gap of unknown length  
\* 11246  
11246: contig of 1284 bp in length  
\* 11326  
11326: gap of unknown length  
\* 11327  
12802: contig of 1476 bp in length  
\* 12803  
12882: gap of unknown length  
\* 12883  
14230: contig of 1348 bp in length  
\* 14231  
14310: gap of unknown length  
\* 14311  
15711: contig of 1401 bp in length



```
* 15712 15791: gap of unknown length
* 15792 16237: contig of 446 bp in length
* 16238 16317: gap of unknown length
* 16318 17991: contig of 1674 bp in length
* 17992 18071: gap of unknown length
* 18072 19191: contig of 1120 bp in length
* 19192 19271: gap of unknown length
* 19272 21324: contig of 2053 bp in length
* 21325 21404: gap of unknown length
* 21405 23898: contig of 2494 bp in length
* 23899 23979: gap of unknown length
* 23980 27079: contig of 3101 bp in length
* 27080 27160: gap of unknown length
* 27161 29602: contig of 2443 bp in length
* 29603 29682: gap of unknown length
* 29683 31919: contig of 2237 bp in length
* 31920 31999: gap of unknown length
* 32000 34542: contig of 2543 bp in length
* 34543 34622: gap of unknown length
* 34623 36445: contig of 1823 bp in length
* 36446 36525: gap of unknown length
* 36526 40018: contig of 3493 bp in length
* 40019 40098: gap of unknown length
* 40099 42061: contig of 1963 bp in length
* 42062 42141: gap of unknown length
* 42142 48164: contig of 6023 bp in length
* 48165 48244: gap of unknown length
* 48245 54296: contig of 6052 bp in length
* 54297 54376: gap of unknown length
* 54377 66627: contig of 12251 bp in length
* 66628 66707: gap of unknown length
* 66708 77734: contig of 11027 bp in length
* 77735 77814: gap of unknown length
* 77815 88534: contig of 10720 bp in length
* 88535 88614: gap of unknown length
* 88615 101020: contig of 12406 bp in length
* 101021 101100: gap of unknown length
* 101101 115583: contig of 14483 bp in length
* 115584 116236: contig of 573 bp in length
* 116237 116317: gap of unknown length
* 116318 116809: contig of 493 bp in length
* 116810 116889: gap of unknown length
* 116890 117344: contig of 454 bp in length
* 117345 117424: gap of unknown length
* 117425 118337: contig of 914 bp in length
* 118338 118417: gap of unknown length
* 118418 118910: contig of 493 bp in length
* 118911 118990: gap of unknown length
* 118991 119603: contig of 613 bp in length
* 119604 119684: gap of unknown length
* 119685 120333: contig of 650 bp in length
* 120334 120413: gap of unknown length
* 120414 120968: contig of 555 bp in length
* 120969 121049: gap of unknown length
* 121050 121686: gap of unknown length
* 121687 122392: contig of 614 bp in length
* 122393 122459: gap of unknown length
* 122460 122949: contig of 490 bp in length
* 122950 123030: gap of unknown length
* 123031 123632: contig of 603 bp in length
* 123633 123712: gap of unknown length
* 123713 124146: contig of 434 bp in length
* 124147 124266: gap of unknown length
* 124267 124732: contig of 506 bp in length
* 124733 124812: gap of unknown length
* 124813 125420: contig of 608 bp in length
* 125421 125500: gap of unknown length
* 125501 126104: contig of 604 bp in length
* 126105 126184: gap of unknown length
* 126185 126663: contig of 479 bp in length
* 126664 126743: gap of unknown length

* 126744 127183: contig of 440 bp in length
* 127184 127863: gap of unknown length
* 127864 127884: contig of 540 bp in length
* 127885 127886: gap of unknown length
* 127887 127888: contig of 842 bp in length
* 127889 128805: gap of unknown length
* 128806 129167: contig of 362 bp in length
* 129168 129247: gap of unknown length
* 129248 129335: contig of 688 bp in length
* 129336 130015: gap of unknown length
* 130016 130560: contig of 545 bp in length
* 130561 130640: gap of unknown length
* 130641 131152: contig of 512 bp in length
* 131153 131232: gap of unknown length
* 131233 131826: contig of 594 bp in length
* 131827 132806: gap of unknown length
* 132807 132884: contig of 477 bp in length
* 132885 133016: contig of 553 bp in length
* 133017 133096: gap of unknown length
* 133097 133747: contig of 651 bp in length
* 133748 133827: gap of unknown length
* 133828 134276: contig of 449 bp in length
* 134277 134356: gap of unknown length
* 134357 134964: contig of 608 bp in length
* 134965 135044: gap of unknown length
* 135045 135383: contig of 339 bp in length
* 135384 135463: gap of unknown length
* 135464 136015: contig of 552 bp in length
* 136016 136095: gap of unknown length
* 136096 136547: contig of 452 bp in length
* 136548 136627: gap of unknown length
* 136628 137067: contig of 440 bp in length
* 137068 137147: gap of unknown length
* 137148 137600: contig of 453 bp in length
* 137601 137680: gap of unknown length
* 137681 138054: contig of 374 bp in length
* 138055 138134: gap of unknown length
* 138135 138352: contig of 218 bp in length
* 138353 138432: gap of unknown length
* 138433 138859: contig of 427 bp in length
* 138860 138939: gap of unknown length
* 138940 139149: contig of 210 bp in length
* 139150 139229: gap of unknown length
* 139230 139727: contig of 498 bp in length
* 139728 139807: gap of unknown length

alignment_scores:
Quality: 756.50 Length: 197
Ratio: 4.274 Gaps: 3
Percent Similarity: 89.848 Percent Identity: 71.574

alignment_block:
US-09-525-867-1 x AC010919 ..
Align seg 1/1 to: AC010919 from: 1 to: 140685

18 ArgSerSerValGlyProAlaValGlnAlaArgGlyValHisGlnSerVa 34
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
103611 CGCTGTGAATTGGCCCTGGT.....GCGCCACAAACAACCTCCAG 103654

34 lAlaThrAspGlyProSerSerThrGlnProAlaLeuProLysAlaArgA 51
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
103655 TTGCCGAAGTGGCCCAAAATCTGCCAAGAGGGCTA.....CTCT 103695

51 laVal.lAlaProLysProSerSerArgGlyGluTyrValValAlaLysLe 67
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
103696 CCGTTGCGCACCAACACAGTCTCGGTTTCGCGAGTGGTCACTGCCACAG 103745

67 uAspAspLeuValAsnTrpAlaArgArgSerSerLeuTrpProMetThrP 84
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
103746 GGACGATCTGCTCAACTGGGTGCGAAGGGCTCGATCGCCACTGACTT 103795
```





59157 GTTAAAGGATGAAGACGCTCCAGATGTGGTATAGGAAG 59195  
seq\_name: gb\_in:AE003500

seq\_documentation\_block:  
LOCUS AE003500 327446 bp DNA INV 04-OCT-2000  
DEFINITION Drosophila melanogaster genomic scaffold 14200001386053 section 17  
of 30 complete sequence.  
ACCESSION AE003500 AE002593  
VERSION AE003500.2 GI:10728273  
KEYWORDS HTG.  
SOURCE fruit fly.  
ORGANISM Drosophila melanogaster  
Eukaryota: Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;  
Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.  
1 (bases 1 to 327446)  
Adams M.D., Celniker, S.E., Holt, R.A., Evans, C.A., Gocayne, J.D.,  
Amaratunga, P.G., Scherer, S.E., Li, P.W., Hoskins, R.A., Galle, R.F.,  
George, R.A., Lewis, S.E., Richards, S., Ashburner, M., Henderson, S.N.,  
Sutton, G.G., Wortman, J.R., Yandell, M.D., Zhang, O., Chen, L.X.,  
Brandon, R.C., Rogers, Y.H., Blazej, R.G., Champe, M., Pfeiffer, B.D.,  
Wap, K.H., Doyle, C., Baxter, E.G., Helt, G., Nelson, C.R., Gabor  
Milos, G.L., Abril, J.F., Agayani, A., An, H.J.,  
Andrews-Franknoch, C., Baldwin, D., Ballew, R.M., Basu, A.,  
Baxendale, J., Bayraktaroglu, L., Beasley, E.M., Beeson, K.Y.,  
Benos, P.V., Berman, B.P., Bhandari, D., Bolshakov, S., Borkova, D.,  
Botchan, M.R., Bouck, J., Brokstein, P., Brottier, P., Burtis, K.C.,  
Busam, D.A., Butler, H., Cadieu, E., Center, A., Chandra, I.,  
Cherry, J.M., Cawley, S., Dahlke, C., Davenport, L.B., Davies, P., de  
Pablos, B., Delcher, A., Deng, Z., Mays, A.D., Dew, I., Dietz, S.M.,  
Doudon, K., Doup, L.E., Downes, M., Dugan-Rocha, S., Dunkov, B.C.,  
Dunn, P., Durbin, K.J., Evangelista, C.C., Ferraz, C., Ferreira, S.,  
Fleischmann, W., Fowler, C., Gabriellian, A.E., Garg, N.S.,  
Gelbart, W.M., Glasser, K., Glodek, A., Gong, F., Gorrell, J.H., Gu, Z.,  
Guan, P., Harris, M., Harris, N.L., Harvey, D., Heiman, T.J.,  
Hernandez, J.R., Houch, J., Hostin, D., Houston, K.A., Howland, T.J.,  
Wei, M.H., Ibegwam, C., Jaitani, M., Kalush, F., Karpen, G.H., Ke, B.,  
Kennison, J.A., Ketchum, K.A., Kimmel, B.E., Kodira, C.D., Kraft, C.,  
Kravitz, S., Kulp, D., Lai, Z., Lasko, P., Lei, Y., Levitsky, A.A.,  
Li, J., Li, Z., Liang, Y., Lin, X., Liu, X., Mattel, B., McIntosh, T.C.,  
McLeod, M.P., McPherson, D., Merkulov, G., Milshina, N.V., Mobarry, C.,  
Morris, J., Moshrefi, A., Mount, S.M., Moy, M., Murphy, B., Murphy, L.,  
Muzny, D.M., Nelson, D.L., Nelson, D.R., Nelson, K.A., Nixon, K.,  
Nuskern, D.R., Paclet, J.M., Palazzolo, M., Pittman, G.S., Pan, S.,  
Pollard, J., Puri, V., Reese, M.G., Reinert, K., Remington, K.,  
Saunders, R.D., Scheeler, F., Shen, H., Shue, B.C., Siden-Kiamos, I.,  
Simpson, M., Skupski, M.P., Smith, T., Spier, E., Spradling, A.C.,  
Stapleton, M., Strong, R., Sun, E., Svirska, R., Tector, C., Turner, R.,  
Venter, E., Wang, A.H., Wang, X., Wang, Z.Y., Wasserman, D.A.,  
Weinstock, G.M., Weissbach, J., Williams, S.M., Woodage, T.,  
Worley, K.C., Wu, D., Yang, S., Yao, Q.A., Ye, J., Yen, R.F.,  
Zaveri, J.S., Zhang, W., Zhou, X., Zhu, S., Zhu, X., Smith, H.O.,  
Zhong, F.N., Zhong, W., Zhou, X., Zhu, S., Zhu, X., Smith, H.O.,  
Gibbs, R.A., Myers, E.W., Rubin, G.M. and Venter, J.C.  
The genome sequence of Drosophila melanogaster  
Science 287 (5461), 2185-2195 (2000)  
20196006  
2 (bases 1 to 327446)  
Adams M.D., Celniker, S.E., Gibbs, R.A., Rubin, G.M. and Venter, J.C.  
Direct Submission  
Submitted (21-MAR-2000) Celera Genomics, 45 West Gude Drive,  
Rockville, MD, USA  
On Oct 9, 2000 this sequence version replaced gi:7293108.  
Location/Qualifiers  
1. 327446  
/organism="Drosophila melanogaster"  
/db\_xref="taxon:7227"  
/chromosome="X"  
complement(join(772, 903,968, 1504,1572, 2077,2153, 2821,  
4145, >4380))  
/gene="CG9281"  
/product="CT26402"

gene  
/db\_xref="FLYBASE:FBan0009281"  
/db\_xref="FLYBASE:FBgn0030672"  
complement(<772, >4380)  
/gene="CG9281"  
/map="13E12-13E14"  
/db\_xref="FLYBASE:FBan0009281"  
/db\_xref="FLYBASE:FBgn0030672"  
complement(join(772, 903,968, 1504,1572, 2077,2153, 2813))  
/gene="CG9281"  
/note="CG9281 gene product"  
/codon\_start=1  
/db\_xref="FLYBASE:FBan0009281"  
/db\_xref="FLYBASE:FBgn0030672"  
/protein\_id="AAP48493.1"  
/db\_xref="GI:7293109"  
/translation="MSSRAKKREAQKKDKARNKKIQIPSTKKANCVPERELTEBEKL  
CAKLEEARISAEARSCTSLAVHPSRDVYKIANFSITFFGSELLQDMLENCGRY  
GLIANGCGKSSLEDAQEIQIDYERLDDMSADIAEVAARILHGLGFKAMQKQAK  
EKLAELAESEEDDAEQEIQIDYERLDDMSADIAEVAARILHGLGFKAMQKQAK  
FSGWMRIARALAEVFKPHLLGLLEDEPTNHLDDACVLEELKTYKRIILSHISQD  
FLNGVCTNIHLTKGLKYTYGYEAFVTRMELLENQMKQYNWQDQISHMKNIAR  
FSGSKALARQAOSKEKTLAKVAQGLTEKVSDDKVLNFFYFPCGQVPPVIMQVNSH  
FRYDEPPIYKLEIGLIDLDPELALVGPNGAGKSTLLKLLXGDLVPTSGMIRKNVNL  
RIARYHGLHELLDLDASPLEYMMRAFDPYKEENMKIIGRYGLIGQQQCPISQLS  
DQGRVYFAWLAWQPHLLLDLEPTNHLDMETIDALADINDFDGGMVLSHDFALI  
NOVAEEIWOCEKETYTKWGGIDYDKHLKNKITSNERKAKAGK"  
join(<4963, 5295,5348, >5848)  
/gene="CG15601"  
/product="CT35724"  
/db\_xref="FLYBASE:FBan0015601"  
/db\_xref="FLYBASE:FBgn0030673"  
/evidence=not\_experimental  
<4963, >5848  
/gene="CG15601"  
/map="13E14-13E14"  
/db\_xref="FLYBASE:FBan0015601"  
/db\_xref="FLYBASE:FBgn0030673"  
/evidence=not\_experimental  
join(4963, 5295,5348, 5848)  
/gene="CG15601"  
/note="CG15601 gene product"  
/codon\_start=1  
/db\_xref="FLYBASE:FBan0015601"  
/db\_xref="FLYBASE:FBgn0030673"  
/evidence=not\_experimental  
/protein\_id="AAF48494.1"  
/db\_xref="GI:7293110"  
/translation="MSKLSDTKIPTEFLEAYRRQPCLYNTLDSYKNRVSREAYGA  
IIRSLKPOLTVSDIKLIKSVTVYKELRWREKELGRTYEPKLPFLRLADGFLR  
SVLSHCKRGKKNSSAQLTIKSDTSKLTCLTAADITMSDALDEEAEAVNGEPE  
ECPLESRPTASICKDSTLCLADQPOQEHYSGQSSQQLPHTMAQKSKYITSLDS  
AGEDDLIFQGSIAQSQRITLPPDSYSAKLRIQVFEAFEGQGFSTEVNSTQLQNT  
F"  
join(<7954, 8220,8556, 8795,8865, 9086,9145, 9416,  
9497, 9993,10131, 10515,10567, 15238,15758, 18218,  
18282, 18755,18814, 19154,19217, 20287,20378, 20513,  
20573, 22032,22060, 23197,23270, 23671,23739, 23927,  
23992, 24313,24377, 24542,24617, 24807,24885, >24987)  
/gene="CG8184"  
/product="CT24358"  
/db\_xref="FLYBASE:FBan0008184"  
/db\_xref="FLYBASE:FBgn0030674"  
<7954, >24987  
/gene="CG8184"  
/map="13E14-13E16"  
/db\_xref="FLYBASE:FBan0008184"  
/db\_xref="FLYBASE:FBgn0030674"  
join(7954, 8220,8556, 8795,8865, 9086,9145, 9416,  
9497, 9993,10131, 10515,10567, 15238,15758, 18218,  
18282, 18755,18814, 19154,19217, 20287,20378, 20513,  
20573, 22032,22060, 23197,23270, 23671,23739, 23927,  
23992, 24313,24377, 24542,24617, 24807,24885, >24987)  
/gene="CG8184"

gene  
/db\_xref="FLYBASE:FBan0008184"  
/db\_xref="FLYBASE:FBgn0030674"  
<7954, >24987  
/gene="CG8184"  
/map="13E14-13E16"  
/db\_xref="FLYBASE:FBan0008184"  
/db\_xref="FLYBASE:FBgn0030674"  
join(7954, 8220,8556, 8795,8865, 9086,9145, 9416,  
9497, 9993,10131, 10515,10567, 15238,15758, 18218,  
18282, 18755,18814, 19154,19217, 20287,20378, 20513,  
20573, 22032,22060, 23197,23270, 23671,23739, 23927,  
23992, 24313,24377, 24542,24617, 24807,24885, >24987)  
/gene="CG8184"

CDS  
/db\_xref="FLYBASE:FBan0008184"  
/db\_xref="FLYBASE:FBgn0030674"  
<7954, >24987  
/gene="CG8184"  
/map="13E14-13E16"  
/db\_xref="FLYBASE:FBan0008184"  
/db\_xref="FLYBASE:FBgn0030674"  
join(7954, 8220,8556, 8795,8865, 9086,9145, 9416,  
9497, 9993,10131, 10515,10567, 15238,15758, 18218,  
18282, 18755,18814, 19154,19217, 20287,20378, 20513,  
20573, 22032,22060, 23197,23270, 23671,23739, 23927,  
23992, 24313,24377, 24542,24617, 24807,24885, >24987)  
/gene="CG8184"

gene  
/db\_xref="FLYBASE:FBan0008184"  
/db\_xref="FLYBASE:FBgn0030674"  
<7954, >24987  
/gene="CG8184"  
/map="13E14-13E16"  
/db\_xref="FLYBASE:FBan0008184"  
/db\_xref="FLYBASE:FBgn0030674"  
join(7954, 8220,8556, 8795,8865, 9086,9145, 9416,  
9497, 9993,10131, 10515,10567, 15238,15758, 18218,  
18282, 18755,18814, 19154,19217, 20287,20378, 20513,  
20573, 22032,22060, 23197,23270, 23671,23739, 23927,  
23992, 24313,24377, 24542,24617, 24807,24885, >24987)  
/gene="CG8184"

CDS  
/db\_xref="FLYBASE:FBan0008184"  
/db\_xref="FLYBASE:FBgn0030674"  
<7954, >24987  
/gene="CG8184"  
/map="13E14-13E16"  
/db\_xref="FLYBASE:FBan0008184"  
/db\_xref="FLYBASE:FBgn0030674"  
join(7954, 8220,8556, 8795,8865, 9086,9145, 9416,  
9497, 9993,10131, 10515,10567, 15238,15758, 18218,  
18282, 18755,18814, 19154,19217, 20287,20378, 20513,  
20573, 22032,22060, 23197,23270, 23671,23739, 23927,  
23992, 24313,24377, 24542,24617, 24807,24885, >24987)  
/gene="CG8184"

gene  
/db\_xref="FLYBASE:FBan0008184"  
/db\_xref="FLYBASE:FBgn0030674"  
<7954, >24987  
/gene="CG8184"  
/map="13E14-13E16"  
/db\_xref="FLYBASE:FBan0008184"  
/db\_xref="FLYBASE:FBgn0030674"  
join(7954, 8220,8556, 8795,8865, 9086,9145, 9416,  
9497, 9993,10131, 10515,10567, 15238,15758, 18218,  
18282, 18755,18814, 19154,19217, 20287,20378, 20513,  
20573, 22032,22060, 23197,23270, 23671,23739, 23927,  
23992, 24313,24377, 24542,24617, 24807,24885, >24987)  
/gene="CG8184"

CDS  
/db\_xref="FLYBASE:FBan0008184"  
/db\_xref="FLYBASE:FBgn0030674"  
<7954, >24987  
/gene="CG8184"  
/map="13E14-13E16"  
/db\_xref="FLYBASE:FBan0008184"  
/db\_xref="FLYBASE:FBgn0030674"  
join(7954, 8220,8556, 8795,8865, 9086,9145, 9416,  
9497, 9993,10131, 10515,10567, 15238,15758, 18218,  
18282, 18755,18814, 19154,19217, 20287,20378, 20513,  
20573, 22032,22060, 23197,23270, 23671,23739, 23927,  
23992, 24313,24377, 24542,24617, 24807,24885, >24987)  
/gene="CG8184"

gene  
/db\_xref="FLYBASE:FBan0008184"  
/db\_xref="FLYBASE:FBgn0030674"  
<7954, >24987  
/gene="CG8184"  
/map="13E14-13E16"  
/db\_xref="FLYBASE:FBan0008184"  
/db\_xref="FLYBASE:FBgn0030674"  
join(7954, 8220,8556, 8795,8865, 9086,9145, 9416,  
9497, 9993,10131, 10515,10567, 15238,15758, 18218,  
18282, 18755,18814, 19154,19217, 20287,20378, 20513,  
20573, 22032,22060, 23197,23270, 23671,23739, 23927,  
23992, 24313,24377, 24542,24617, 24807,24885, >24987)  
/gene="CG8184"

CDS  
/db\_xref="FLYBASE:FBan0008184"  
/db\_xref="FLYBASE:FBgn0030674"  
<7954, >24987  
/gene="CG8184"  
/map="13E14-13E16"  
/db\_xref="FLYBASE:FBan0008184"  
/db\_xref="FLYBASE:FBgn0030674"  
join(7954, 8220,8556, 8795,8865, 9086,9145, 9416,  
9497, 9993,10131, 10515,10567, 15238,15758, 18218,  
18282, 18755,18814, 19154,19217, 20287,20378, 20513,  
20573, 22032,22060, 23197,23270, 23671,23739, 23927,  
23992, 24313,24377, 24542,24617, 24807,24885, >24987)  
/gene="CG8184"

gene  
/db\_xref="FLYBASE:FBan0008184"  
/db\_xref="FLYBASE:FBgn0030674"  
<7954, >24987  
/gene="CG8184"  
/map="13E14-13E16"  
/db\_xref="FLYBASE:FBan0008184"  
/db\_xref="FLYBASE:FBgn0030674"  
join(7954, 8220,8556, 8795,8865, 9086,9145, 9416,  
9497, 9993,10131, 10515,10567, 15238,15758, 18218,  
18282, 18755,18814, 19154,19217, 20287,20378, 20513,  
20573, 22032,22060, 23197,23270, 23671,23739, 23927,  
23992, 24313,24377, 24542,24617, 24807,24885, >24987)  
/gene="CG8184"

CDS  
/db\_xref="FLYBASE:FBan0008184"  
/db\_xref="FLYBASE:FBgn0030674"  
<7954, >24987  
/gene="CG8184"  
/map="13E14-13E16"  
/db\_xref="FLYBASE:FBan0008184"  
/db\_xref="FLYBASE:FBgn0030674"  
join(7954, 8220,8556, 8795,8865, 9086,9145, 9416,  
9497, 9993,10131, 10515,10567, 15238,15758, 18218,  
18282, 18755,18814, 19154,19217, 20287,20378, 20513,  
20573, 22032,22060, 23197,23270, 23671,23739, 23927,  
23992, 24313,24377, 24542,24617, 24807,24885, >24987)  
/gene="CG8184"

gene  
/db\_xref="FLYBASE:FBan0008184"  
/db\_xref="FLYBASE:FBgn0030674"  
<7954, >24987  
/gene="CG8184"  
/map="13E14-13E16"  
/db\_xref="FLYBASE:FBan0008184"  
/db\_xref="FLYBASE:FBgn0030674"  
join(7954, 8220,8556, 8795,8865, 9086,9145, 9416,  
9497, 9993,10131, 10515,10567, 15238,15758, 18218,  
18282, 18755,18814, 19154,19217, 20287,20378, 20513,  
20573, 22032,22060, 23197,23270, 23671,23739, 23927,  
23992, 24313,24377, 24542,24617, 24807,24885, >24987)  
/gene="CG8184"

CDS  
/db\_xref="FLYBASE:FBan0008184"  
/db\_xref="FLYBASE:FBgn0030674"  
<7954, >24987  
/gene="CG8184"  
/map="13E14-13E16"  
/db\_xref="FLYBASE:FBan0008184"  
/db\_xref="FLYBASE:FBgn0030674"  
join(7954, 8220,8556, 8795,8865, 9086,9145, 9416,  
9497, 9993,10131, 10515,10567, 15238,15758, 18218,  
18282, 18755,18814, 19154,19217, 20287,20378, 20513,  
20573, 22032,22060, 23197,23270, 23671,23739, 23927,  
23992, 24313,24377, 24542,24617, 24807,24885, >24987)  
/gene="CG8184"

gene  
/db\_xref="FLYBASE:FBan0008184"  
/db\_xref="FLYBASE:FBgn0030674"  
<7954, >24987  
/gene="CG8184"  
/map="13E14-13E16"  
/db\_xref="FLYBASE:FBan0008184"  
/db\_xref="FLYBASE:FBgn0030674"  
join(7954, 8220,8556, 8795,8865, 9086,9145, 9416,  
9497, 9993,10131, 10515,10567, 15238,15758, 18218,  
18282, 18755,18814, 19154,19217, 20287,20378, 20513,  
20573, 22032,22060, 23197,23270, 23671,23739, 23927,  
23992, 24313,24377, 24542,24617, 24807,24885, >24987)  
/gene="CG8184"

CDS  
/db\_xref="FLYBASE:FBan0008184"  
/db\_xref="FLYBASE:FBgn0030674"  
<7954, >24987  
/gene="CG8184"  
/map="13E14-13E16"  
/db\_xref="FLYBASE:FBan0008184"  
/db\_xref="FLYBASE:FBgn0030674"  
join(7954, 8220,8556, 8795,8865, 9086,9145, 9416,  
9497, 9993,10131, 10515,10567, 15238,15758, 18218,  
18282, 18755,18814, 19154,19217, 20287,20378, 20513,  
20573, 22032,22060, 23197,23270, 23671,23739, 23927,  
23992, 24313,24377, 24542,24617, 24807,24885, >24987)  
/gene="CG8184"

gene  
/db\_xref="FLYBASE:FBan0008184"  
/db\_xref="FLYBASE:FBgn0030674"  
<7954, >24987  
/gene="CG8184"  
/map="13E14-13E16"  
/db\_xref="FLYBASE:FBan0008184"  
/db\_xref="FLYBASE:FBgn0030674"  
join(7954, 8220,8556, 8795,8865, 9086,9145, 9416,  
9497, 9993,10131, 10515,10567, 15238,15758, 18218,  
18282, 18755,18814, 19154,19217, 20287,20378, 20513,  
20573, 22032,22060, 23197,23270, 23671,23739, 23927,  
23992, 24313,24377, 24542,24617, 24807,24885, >24987)  
/gene="CG8184"

CDS  
/db\_xref="FLYBASE:FBan0008184"  
/db\_xref="FLYBASE:FBgn0030674"  
<7954, >24987  
/gene="CG8184"  
/map="13E14-13E16"  
/db\_xref="FLYBASE:FBan0008184"  
/db\_xref="FLYBASE:FBgn0030674"  
join(7954, 8220,8556, 8795,8865, 9086,9145, 9416,  
9497, 9993,10131, 10515,10567, 15238,15758, 18218,  
18282, 18755,18814, 19154,19217, 20287,20378, 20513,  
20573, 22032,22060, 23197,23270, 23671,23739, 23927,  
23992, 24313,24377, 24542,24617, 24807,24885, >24987)  
/gene="CG8184"

gene  
/db\_xref="FLYBASE:FBan0008184"  
/db\_xref="FLYBASE:FBgn0030674"  
<7954, >24987  
/gene="CG8184"  
/map="13E14-13E16"  
/db\_xref="FLYBASE:FBan0008184"  
/db\_xref="FLYBASE:FBgn0030674"  
join(7954, 8220,8556, 8795,8865, 9086,9145, 9416,  
9497, 9993,10131, 10515,10567, 15238,15758, 18218,  
18282, 18755,18814, 19154,19217, 20287,20378, 20513,  
20573, 22032,22060, 23197,23270, 23671,23739, 23927,  
23992, 24313,24377, 24542,24617, 24807,24885, >24987)  
/gene="CG8184"

CDS  
/db\_xref="FLYBASE:FBan0008184"  
/db\_xref="FLYBASE:FBgn0030674"  
<7954, >24987  
/gene="CG8184"  
/map="13E14-13E16"  
/db\_xref="FLYBASE:FBan0008184"  
/db\_xref="FLYBASE:FBgn0030674"  
join(7954, 8220,8556, 8795,8865, 9086,9145, 9416,  
9497, 9993,10131, 10515,10567, 15238,15758, 18218,  
18282, 18755,18814, 19154,19217, 20287,20378, 20513,  
20573, 22032,22060, 23197,23270, 23671,23739, 23927,  
23992, 24313,24377, 24542,24617, 24807,24885, >24987)  
/gene="CG8184"

gene  
/db\_xref="FLYBASE:FBan0008184"  
/db\_xref="FLYBASE:FBgn0030674"  
<7954, >24987  
/gene="CG8184"  
/map="13E14-13E16"  
/db\_xref="FLYBASE:FBan0008184"  
/db\_xref="FLYBASE:FBgn0030674"  
join(7954, 8220,8556, 8795,8865, 9086,9145, 9416,  
9497, 9993,10131, 10515,10567, 15238,15758, 18218,  
18282, 18755,18814, 19154,19217, 20287,20378, 20513,  
20573, 22032,22060, 23197,23270, 23671,23739, 23927,  
23992, 24313,24377, 24542,24617, 24807,24885, >24987)  
/gene="CG8184"

CDS  
/db\_xref="FLYBASE:FBan0008184"  
/db\_xref="FLYBASE:FBgn0030674"  
<7954, >24987  
/gene="CG8184"  
/map="13E14-13E16"  
/db\_xref="FLYBASE:FBan0008184"  
/db\_xref="FLYBASE:FBgn0030674"  
join(7954, 8220,8556, 8795,8865, 9086,9145, 9416,  
9497, 9993,10131, 10515,10567, 15238,15758, 18218,  
18282, 18755,18814, 19154,19217, 20287,20378, 20513,  
20573, 22032,22060, 23197,23270, 23671,23739, 23927,  
23992, 24313,24377, 24542,24617, 24807,24885, >24987)  
/gene="CG8184"

gene  
/db\_xref="FLYBASE:FBan0008184"  
/db\_xref="FLYBASE:FBgn0030674"  
<7954, >24987  
/gene="CG8184"  
/map="13E14-13E16"  
/db\_xref="FLYBASE:FBan0008184"  
/db\_xref="FLYBASE:FBgn0030674"  
join(7954, 8220,8556, 8795,8865, 9086,9145, 9416,  
9497, 9993,10131, 10515,10567, 15238,15758, 18218,  
18282, 18755,18814, 19154,19217, 20287,20378, 20513,  
20573, 22032,22060, 23197,23270, 23671,23739, 23927,  
23992, 24313,24377, 24542,24617, 24807,24885, >24987)  
/gene="CG8184"

CDS  
/db\_xref="FLYBASE:FBan0008184"  
/db\_xref



```

BASE COUNT      23093 a 18268 c 19032 g 23214 t
ORIGIN
alignment_scores:
  Quality: 743.00      Length: 154
  Ratio: 4.921         Gaps: 0
Percent Similarity: 98.052 Percent Identity: 83.766
alignment_block:
US-09-525-867-1 x AC019889/rev ..
Align seg 1/1 to reverse of: AC019889 from: 1 to: 83607
60 GlyGluTyrValValAlaLysLeuAspLeuValAsnTrpAlaArgAr 76
14234 GCGAGTGGACCTGCGCGAGTGGATGACCTGCTAACTGGGACGCAA 14185
76 gSerSerLeuTrpProMetThrPheGlyLeuAlaCysCysAlaValGlu 93
14184 GGGATCCCTGCGCGCGTGGCTTGGTCTGCGCTGCTGCGCGCTCGAAA 14135
93 etMetHisMetalAlaProArgTyrAspMetAspArgPheGlyValVal 109
14134 TGATCAGATTCGCGCTCTCGCTACGACATGGATCGCTACGGTGTGGT 14085
110 PheArgAlaSerProArgGlnSerAspValMetLeuValAlaGlyThrLe 126
14084 TTCGAGCATCTCCCGCAGCGGACGTGCTCATCGTGGCCGGAACCT 14035
126 uThrAsnLysMetalAlaProAlaLeuArgLysValTyrAspGlnMetPro 143
14034 GACCAACAGATGGCAGCGCTTCGGAAGATCTACGACCATGCGCG 13985
143 LuProArgTyrValValSerMetGlySerCysAlaAsnGlyGlyTyr 159
13984 AGCCGAGATGGTCAATTCATGCGGAGTGGCAATGTTGGCGGTAC 13935
160 TyrHisTyrSerTyrSerValValArgGlyCysAspArgLeuValProVa 176
13934 TACCACTACTCTACTCGTGGTTCGCGGCTGGCATGTCATGTCGGT 13885
176 LaspleTyrIleProGlyCysProProThrAlaGluAlaLeuTyrG 193
13884 GGACATCTACGTGCGCGATGTCGCCACCGCGAGGCTTAATGTAGC 13835
193 LylleuGlnLeuGlnArgLysLysLysArgGluArgArgLeuGluIle 209
13834 GAATCTGAGCTGAGAGGAGGTAAGGCTGAGGACCCCTGCAGATG 13785
210 TrpTyrArgArg 213
13784 TGTACAGGAAG 13773
seq_name: gb_in:AC007856
seq_documentation_block:
LOCUS      AC007856.1 175124 bp DNA INV 17-FEB-2001
DEFINITION Drosophila melanogaster, chromosome 3R, region 99B-99B, BAC clone
ACCESSION  BACR05L03, complete sequence.
VERSION    AC007856
KEYWORDS   HTG.
SOURCE     fruit fly.
ORGANISM   Drosophila melanogaster
            Eukaryota; Neozoa; Arthropoda; Tracheata; Hexapoda; Insecta;
            Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
            Muscomorpha; Ephyraoidea; Drosophilidae; Drosophila.
REFERENCE  1 (bases 1 to 175124)
            Celisner,S.E., Adams,M.D., Kronmiller,B., Tyler,D., Wan,K.H.,
            Holt,R.A., Evans,C.A., Gocayne,J.D., Amanatides,P.G., Brandon,R.C.,
            Rogers,Y., An,H., Baldwin,D., Banzon,J., Beeson,K.Y., Busam,D.A.,

```

```

Carlson,J.W., Center,A., Champe,M., Davenport,L.B., Dietz,S.M.,
Dodson,K., Dorsett,V., Doup,L.E., Doyle,C., Dresnek,D., Farfan,D.,
Ferreira,S., Frise,E., Galle,R.F., Garg,N.S., George,R.A.,
Gonzalez,M., Houck,J., Hoskins,R.A., Hostin,D., Howland,T.J.,
Ibegwam,C., Jalali,M., Kruse,D., Li,P., Mattei,B., Moshrefi,A.,
McIntosh,T.C., Moy,M., Murphy,B., Nelson,C., Nelson,K.A., Nunoo,J.,
Pacleb,J., Paragas,V., Park,S., Patel,S., Pfeiffer,B., Scheeler,F.,
Phouanavong,S., Pittman,G.S., Puri,V., Richards,S., Scheeler,F.,
Stapleton,M., Strong,R., Svirskas,R., Tector,C., Williams,S.M.,
Zaveri,J.S., Smith,H.O., Rubin,G.M. and Venter,J.C.
Sequencing of Drosophila chromosome 3R, region 99B-99B
Unpublished
2 (bases 1 to 175124)
Celisner,S.E., Agbayani,A., Arcaina,T.T., Baxter,E., Blazej,R.G.,
Butenhorf,C., Champe,M., Chavez,C., Chew,M., Ciesiolka,L.,
Doyle,C.M., Farfan,D.E., Galle,R., George,R.A., Harris,N.L.,
Hoskins,R.A., Houston,K.A., Hummasti,S.R., Karra,K., Kearney,L.,
Kim,E., Lee,B., Lewis,S., Li,P., Lomotan,M.A., Mazda,P.,
Moshrefi,A.R., Moshrefi,M., Nixon,K., Pacleb,J.M., Park,S.,
Pfeiffer,B., Poon,L., Sequeira,A., Sethi,H., Snir,E.,
Svirskas,R.R., Wan,K.H., Weinburg,T., Zhang,R., Zieran,L.L. and
Rubin,G.M.
Direct Submission
Submitted (17-JUN-1999) Drosophila Genome Center, Lawrence Berkeley
Laboratory, MS 64-121, Berkeley, CA 94720, USA
On Feb 17, 2001 this sequence version replaced gi:6633905.
Sequence submitted by:
Berkeley Drosophila Genome Project
Lawrence Berkeley National Laboratory, MS 64-121
Berkeley, CA 94720
This sequence was assembled using end sequences from a whole genome
shotgun and from subclones of this BAC and its neighboring clones.
For further information about this sequence, including its location
and relationship to other sequences, please visit our sequence
archive Web site (http://www.fruitfly.org/sequence/) or send email
to bdgpre@fruitfly.berkeley.edu.
FEATURES             source
     Location/Qualifiers
         1..175124
             /organism="Drosophila melanogaster"
             /strain="y; cn bw sp"
             /db_xref="taxon:7227"
             /chromosomes="3R"
             /map="99B-99B"
             /clone="BACR05L03 (D795)"
             /clone_lib="RPCI-98 (Roswell Park Cancer Institute
             Drosophila melanogaster BAC library, partial ECORI in
             pBACe3.6)"
BASE COUNT      47809 a 39206 c 39325 g 48784 t
ORIGIN
alignment_scores:
  Quality: 743.00      Length: 154
  Ratio: 4.921         Gaps: 0
Percent Similarity: 98.052 Percent Identity: 83.766
alignment_block:
US-09-525-867-1 x AC007856
Align seg 1/1 to: AC007856 from: 1 to: 175124
60 GlyGluTyrValValAlaLysLeuAspLeuValAsnTrpAlaArgAr 76
22222 GCGAGTGGACCTGCGCGAGTGGATGACCTGCTAACTGGGACGCAA 22271
76 gSerSerLeuTrpProMetThrPheGlyLeuAlaCysCysAlaValGlu 93
22272 GGGATCCCTGCGCGCTGACCTTGGTCTGCGCTGCTGCGCGCTCGAAA 22321
93 etMetHisMetalAlaProArgTyrAspMetAspArgPheGlyValVal 109
22322 TGATCAGATTCGCGCTCTCGCTACGACATGGATCGCTACGGTGTGGT 22371

```

```

TITLE
JOURNAL
REFERENCE
AUTHORS
COMMENT

```

110 PheArgIaSerProArgGInSerAspValMetIleValAlaGlyThrLe 126  
 22372 TTCGAGCATCTCTCCAGCGGAGCGTCTCATGTCGCGGAAACCT 22421  
 126 uThrAsnLysMetAlaProAlaLeuArgLysValTyrAspGInMetProG 143  
 22422 GACCACAGATGACCGGCCCTTCGGAAGATCTACGACCAAGATCCCG 22471  
 143 luProArgTyrValValSerMetGlySerCysAlaAsnGlyGlyTyr 159  
 22472 AGCCGAGATGGTTCATTTCATGGCAGGTGCGCCCAATGTCGCGCTAC 22521  
 160 TyrHisTyrSerTyrSerValValArgGlyCysAspArgIleValProVa 176  
 22522 TACCACATCTCTACTCGGTCGGTTCGCGGTGCGATGTCGATGTTCCGT 22571  
 176 lAspIleTyrIleProGlyCysProThrAlaGluAlaLeuLeuTyrG 193  
 22572 GGACATCTAGTTCGCGGATGTCGCCGCCAGCGCGAGGCTTAATGATG 22621  
 193 lylLeuGInLeuGInArgLysIleLysArgGluArgLeuGInIle 209  
 22622 GAATCCCTGCAGCTGCAGAGAAGTAAGCGCATGAGGACCGCTGCAGATG 22671  
 210 TrpTyrArgArg 213  
 22672 TGGTACAGGAAG 22683

seq\_name: gb\_in:AE003770

seq\_documentation\_block:  
 LOCUS AE003770 224100 bp DNA INV 05-OCT-2000  
 DEFINITION Drosophila melanogaster genomic scaffold 142000013386035 section 95  
 of 105, complete sequence.  
 ACCESSION AE003770 AE002708  
 VERSION AE003770.1 GI:7301790  
 KEYWORDS fruit fly.  
 SOURCE Drosophila melanogaster  
 ORGANISM Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;  
 Muscomorpha; Ephydroidea; Drosophilidae; Drosophila...

1 (bases 1 to 224100)  
 Adams,M.D., Celniker,S.B., Holt,R.A., Evans,C.A., Gocayne,J.D.,  
 Amanatides,P.G., Scherer,S.E., Li,P.W., Hoskins,R.A., Galie,R.F.,  
 George,R.A., Lewis,S.E., Richards,S., Ashburner,M., Henderson,S.N.,  
 Sutton,G.G., Wortman,J.R., Yandell,M.D., Zhang,Q., Chen,L.X.,  
 Brandon,R.C., Rogers,Y.H., Blazej,R.G., Champe,M., Pfeiffer,B.D.,  
 Wan,K.H., Doyle,C., Baxter,E.G., Helt,G., Nelson,C.R., Gabor  
 Miklos,G.L., Abril,J.F., Agbayani,A., An,H.J., Basu,A.,  
 Andrews-Pfannkoch,C., Baldwin,D., Ballew,R.M., Beeson,K.Y.,  
 Baxendale,J., Bayraktaroglu,L., Beasley,E.M., Beeson,K.Y.,  
 Benos,P.V., Berman,B.P., Bhandari,D., Bolshakov,S., Borkova,D.,  
 Botchan,M.R., Bouck,J., Brokstein,P., Brottier,P., Burtis,K.C.,  
 Busam,D.A., Butler,H., Cadieu,E., Center,A., Chandra,I.,  
 Cherry,J.M., Cawley,S., Dahlke,C., Davenport,L.B., Davies,P., de  
 Pablos,B., Delcher,A., Deng,Z., Mays,A.D., Dew,I., Dietz,S.M.,  
 Dodson,K., Doup,L.E., Downes,M., Dugan-Rocha,S., Dunkov,B.C.,  
 Dunn,P., Durbin,K.J., Evangelista,C.C., Ferraz,C., Ferrieres,S.,  
 Fleischmann,W., Foster,C., Gabrielian,A.E., Garg,N.S.,  
 Gelbart,W.M., Glasser,K., Glodek,A., Gong,F., Gorrell,J.H., Gu,Z.,  
 Guan,P., Harris,M., Harris,N.L., Harvey,D., Heiman,T.J.,  
 Hernandez,J.R., Houck,J., Hostin,D., Houston,K.A., Howland,T.J.,  
 Wei,M.H., Ibegwam,C., Jalali,M., Kalush,F., Karpen,G.H., Ke,Z.,  
 Kenison,J.A., Ketchum,K.A., Kimmel,B.E., Kodira,C.D., Kraft,C.,  
 Kravitz,S., Kulp,D., Lai,Z., Lasko,P., Lei,Y., Levitsky,A.A.,  
 Li,J., Li,Z., Ming,Y., Lin,X., Liu,X., Mattei,B., McIntosh,T.C.,  
 McLeod,M.P., McPherson,D., Merkulov,G., Milshina,N.V., Mobarry,C.,  
 Morris,J., Moshrefi,A., Mount,S.M., Moy,M., Murphy,B., Murphy,L.,  
 Muzny,D.M., Nelson,D.L., Nelson,D.R., Nelson,K.A., Nixon,K.,  
 Nusskern,D.R., Pacle,J.M., Palazzolo,M., Pittman,G.S., Pan,S.,  
 Pohl,J.V., Puri,Y., Reese,M.G., Reinert,K., Remington,K.,  
 Saunders,R.D., Scheeler,F., Shen,H., Shue,B.C., Siden-Kiamos,I.,

Simpson,M., Skupski,M.P., Smith,T., Spier,E., Spradling,A.C.,  
 Stapleton,M., Strong,R., Sun,E., Svirskas,R., Tector,C., Turner,R.,  
 Venter,E., Wang,A.H., Wang,X., Wang,Z.Y., Wassarman,D.A.,  
 Weinstock,G.M., Weissenbach,J., Williams,S.M., Woodage,T.,  
 Worley,K.C., Wu,D., Yang,S., Yao,Q.A., Ye,J., Yeh,R.F.,  
 Zaveri,J.S., Zhan,M., Zhang,G., Zhao,Q., Zheng,L., Zheng,X.H.,  
 Zenger,F.N., Zhong,W., Zhou,X., Zhu,S., Zhu,X., Smith,H.O.,  
 Gibbs,R.A., Myers,E.W., Rubin,G.M. and Venter,J.C.  
 The genome sequence of *Drosophila melanogaster*  
 Science 287 (5461), 2185-2195 (2000)

TITLE  
 JOURNAL  
 MEDLINE  
 REFERENCE  
 AUTHORS  
 TITLE  
 JOURNAL  
 FEATURES  
 source

2 (bases 1 to 224100)  
 Adams,M.D., Celniker,S.E., Gibbs,R.A., Rubin,G.M. and Venter,C.J.  
 Direct Submission  
 Submitted (21-MAR-2000) Celera Genomics, 45 West Gude Drive,  
 Rockville, MD, USA  
 Location/Qualifiers  
 1. 224100  
 /organism="Drosophila melanogaster"  
 /db\_xref="taxon:7227"  
 /chromosome="3R"  
 <6113..>6751  
 /gene="CG2014"  
 /product="CT6471"  
 /db\_xref="FLYBASE:FBan0002014"  
 /db\_xref="FLYBASE:FBgn0039669"  
 <6113..>6751  
 /gene="CG2014"  
 /map="99B3-99B4"  
 /db\_xref="FLYBASE:FBan0002014"  
 /db\_xref="FLYBASE:FBgn0039669"  
 6113..6751  
 /gene="CG2014"  
 /note="CG2014 gene product"  
 /codon\_start=1  
 /db\_xref="FLYBASE:FBan0002014"  
 /db\_xref="FLYBASE:FBgn0039669"  
 /protein\_id="AAF56901.1"  
 /db\_xref="GI:7301791"  
 /translation="MISFLKLARNPGLLLPGSGVLSQSLRGKTPVYDRCQNP  
 KWGSVAFGRNKTGWGTCATDLNLRKGLSLPTFGLACCAVHMHLAARPYDM  
 DRYGVFRASPRQADVLIVAGTLTKMAPRKTYDMPRPVMSGANGSGGYH  
 YSYSVRGCDRIVPVDIVPCPTAEALMGILQLQKKYKVRMTLQWYRK"  
 join(46705..48163,54471..55572)  
 /gene="msh"  
 /note="Nucleotide sequence of the Celera sequence differs  
 from the published sequence for this transcript."  
 /product="CT2801"  
 /db\_xref="FLYBASE:FBan0001897"  
 /db\_xref="FLYBASE:FBgn0002849"  
 <46705..>55572  
 /gene="msh"  
 /note="CG1897"  
 /map="99B4-99B5"  
 /db\_xref="FLYBASE:FBan0001897"  
 /db\_xref="FLYBASE:FBgn0002849"  
 join(47104..48163,54471..54808)  
 /gene="msh"  
 /note="msh gene product; Nucleotide sequence of the Celera  
 sequence differs from the published sequence for this  
 transcript"  
 /codon\_start=1  
 /db\_xref="FLYBASE:FBan0001897"  
 /db\_xref="FLYBASE:FBgn0002849"  
 /protein\_id="AAF56902.1"  
 /db\_xref="GI:7301792"  
 /translation="MSSGNMTSSNLTNLSHPAGLNALASPTSPSALLAHQHLIQ  
 CHQHQHQOQOQQAALQAAVHPHAKHLKTTSLRNSFVSALSLADTPRPQQA  
 ADGPNLTSSAATSPISQASSTPPPPAAAOVPANTFHPAAVHAHLQAAHAA  
 RAHQHQAQALRQOQQAQDARANSPPASTSTPSPIGSAAGSGNVASTPAKNE  
 RHSPGLSTSELYDEEMQDHDHDEEDSDIVEDMADDSPPASTPDGLDGSCK  
 SLESPHGPFGSHMQSTILSPAALASGHVPIRPTFFSALAAAVAWTGMGGVFWPQT  
 RQMPFPGPGFPGAGFGGDANEPPRIKCLRKHKNRKPRTPTTQQLLSLEKPRE

mrna  
KQVLSIAERARSSRLTETQVTFQNRRAKAKRLQAEIEKIKMAALGRGAPGAQ  
WAMAGTFHPSLMLHG  
complement(join(77651..77884,77983..>78412))  
/gene="CG7567"  
/product="Ct8429"  
/db\_xref="FLYBASE:FBan0007567"  
/db\_xref="FLYBASE:FBgn0039670"  
complement(<77651..>78412)  
/gene="CG7567"  
/map="99B5-99B6"  
/db\_xref="FLYBASE:FBan0007567"  
/db\_xref="FLYBASE:FBgn0039670"  
complement(join(77651..77884,77983..78384))  
/gene="CG7567"  
/note="CG7567 gene product"  
/codon\_start=1  
/db\_xref="FLYBASE:FBan0007567"  
/db\_xref="FLYBASE:FBgn0039670"  
/protein\_id="AAF56903.1"  
/db\_xref="GI:7301793"  
translation="MPSKESLLAVLAFGLVSLSLVVKPKSLKLPNTKKILGLQ  
QTELAARDPTSVRCFDIYPIINGLSAQYELDKVCKYDITASELVLAWSLTF  
QTQSDRGNTFFDCSSIVDFVAFCEFANYGAQSKIMYQVSANATEAAVQIKIHL  
QTLDLSQELSNYSEDFEGTSLYGLNKLKLAGAPVQDPTIANWYITTA"  
complement(79354..80123)  
/gene="CG11470"  
/product="Ct36281"  
/db\_xref="FLYBASE:FBan0011470"  
/db\_xref="FLYBASE:FBgn0039671"  
complement(<79354..>80123)  
/gene="CG11470"  
/map="99B5-99B6"  
/db\_xref="FLYBASE:FBan0011470"  
/db\_xref="FLYBASE:FBgn0039671"  
complement(79388..80080)  
/gene="CG11470"  
/note="CG11470 gene product"  
/codon\_start=1  
/db\_xref="FLYBASE:FBan0011470"  
/db\_xref="FLYBASE:FBgn0039671"  
/protein\_id="AAF56904.1"  
/db\_xref="GI:7301794"  
translation="MFAKQAKILVCLSLGLSLAVKLQIPFRPDHAIHQELQSLRELA  
NVCSDTSCFSYIKPLAKIADQFTNETACISAYDNCTSHISEKIAEDKQILIRSA  
EIVCSPTNSQWLTLEQPLPTVWSLECASTNSAKSTFYAISANATIAVQIQE  
QYTLILSRKSGVINDANRSVVEDTSDTYELLNCLNKGPTTTCNPLIYPTITATT  
TTTITWAAPLLR"  
join(82216..82539,83590..84481,84545..84662,87559..87664,  
87737..87791,88482..88621,94828..94899)  
/gene="CG1906"  
/product="Ct2831"  
/db\_xref="FLYBASE:FBan0001906"  
/db\_xref="FLYBASE:FBgn0039672"  
82216..>94899  
/gene="CG1906"  
/map="99B5-99B7"  
/db\_xref="FLYBASE:FBan0001906"  
/db\_xref="FLYBASE:FBgn0039672"  
join(83669..84481,84545..84662,87559..87664,87737..87791,  
88482..88496)  
/gene="CG1906"  
/note="CG1906 gene product"  
/codon\_start=1  
/db\_xref="FLYBASE:FBan0001906"  
/db\_xref="FLYBASE:FBgn0039672"  
/protein\_id="AAF56905.1"  
/db\_xref="GI:7301795"  
translation="MGGFIDKPKAKHNDGEGCNLLFGVSSMQGWRSEMEDAYVARA  
GLGDALDMFFFAVFGHAGCKVSEHCAXHLLSIISTEEFIGGDHVKIGRTGFLRID  
EVNRELPEETRESEKCGGTAVCAVFLQTVVIANGDSRAVLRCQGVPEVATQDKP  
ILPEERERINAGSGVMIKRVNGTILAYSALGDYDFKNVKEGQEQSLVPEPELFCQ  
SQSDIDFLPACGGLWDSNEDVCSFIHSRMVTSNLSIANQVDTCLHKGRDN  
MSIIITIAFGAPAKPTBEAIEAHRLEKQIEKTRDEIESKITTVDYDILLKCLQNRDDI

mrna  
EGLPPGGGLQSKYHVIERTFKQEFDPORPCESEPH"  
complement(join(89615..90119,90172..>91012))  
/gene="CG7568"  
/product="Ct8427"  
/db\_xref="FLYBASE:FBan0007568"  
/db\_xref="FLYBASE:FBgn0039673"  
/evidence=not\_experimental  
complement(<89615..>91012)  
/gene="CG7568"  
/map="99B5-99B7"  
/db\_xref="FLYBASE:FBan0007568"  
/db\_xref="FLYBASE:FBgn0039673"  
alignment\_scores:  
Quality: 743.00 Length: 154  
Ratio: 4.921 Gaps: 0  
Percent Similarity: 98.052 Percent Identity: 83.766  
alignment\_block:  
US-09-525-867-1 x AE003770 ..  
Align seg 1/1 to: AE003770 from: 1 to: 224100  
60 GlyGluTyrValValAlaLysLeuAspLeuValAlaSerPheGlyValVal 76  
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:  
6287 GCGAGTGGACCTGTGCCAGGTGGATGACCTGTAAACTGGGACGCAA 6336  
76 gSerSerLeuTyrProMetThrPheGlyLeuAlaCysCysAlaValGlu 93  
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:  
6337 GGGATCCCTCTGCGCGCTGACCTTTGGTCTGCGCTGCTGCGCGTGGAAA 6386  
93 eMethHisMetAlaAlaProArgTyrAspMetAspArgPheGlyValVal 109  
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:  
6387 TGATGCACATTCGGCTCTCGCTACGACATGGATCGCTACGGTGTGGTG 6436  
110 PheArgAlaSerProArgGlnSerAspValMetIleValAlaGlyThrLe 126  
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:  
6437 TTCCGAGCATCTCTCGCGCGGACGCTGCTCATCGTGGCGGAAACCT 6486  
126 uThrAsnLysMetAlaProAlaLeuArgLysValTyrAspGlnMetProG 143  
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:  
6487 GACCAACRAGATGGCACCGGCTTTTCGGAAGATCTACGCCAGATGCCG 6536  
143 LuProArgTyrValValSerMetGlySerCysAlaAsnGlyClyGlyTyr 159  
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:  
6537 AGCCGAGATGGGTCTATTCATGGCAGTTGGCCCAATGGTGGCGGTAC 6586  
160 TyrHisTyrSerTyrSerValValArgGlyCysAspArgIleValProVa 176  
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:  
6587 TACCACTACTCTCTACTCGTGGTGGTCCGCGCTGCGATCGCATGTTCGGT 6636  
176 LasPletYrIleProGlyCysProProThrAlaGluAlaLeuTyrG 193  
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:  
6637 GGACATCTACGTGCGCGGATGTCCGCCACCGCGGAGGCTTAATGTACG 6686  
193 LyIleLeuGlnLeuGlnArgLysIleLysArgGluArgGluGlnIle 209  
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:  
6687 GAATCTGCGCTGAGAGAGAGGTAAAGCCATGAGGACCTGCAGATG 6736  
210 TriptYrArgArg 213  
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:  
6737 TGGTACAGGAAG 6748  
seq\_name: gb\_in:AC009888  
seq\_documentation\_block:  
LOCUS AC009888 311043 bp DNA INV 04-SEP-1999  
DEFINITION Drosophila melanogaster, chromosome 3R, region 99A7-99B5, BAC  
clones BACR48K01 and BACR48G23, complete sequence.  
ACCESSION AC009888 AC006093 AC007081  
VERSION AC009888.1 GI:5822961  
KEYWORDS fruit fly.



```

ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE
AUTHORS Celniker,S.E., Agbayani,A., Arcaina,T.T., Baxter,E., Blazej,R.G.,
Butenhoff,C., Champe,M., Chavez,C., Chew,M., Ciesiolka,L.,
Doyle,C.M., Farfan,D.E., Galle,R., George,R.A., Harris,N.L.,
Hinkle,A., Hoskins,R.A., Houston,K.A., Hummasti,S.R., Karra,K.,
Kearney,L., Kim,E., Lee,B., Lewis,S., Li,P., Ling,H.,
Moshrefi,A.R., Moshrefi,M., Nixon,K., Pacleb,J.M., Park,S.,
Pfeiffer,B., Poon,L., Richards,S., Sequeira,A., Sethi,H.,
Svirskas,R.R., Wan,K.H., Webster,D., Woolley,P., Yang,S., Yee,M.,
Zieran,L.L. and Rubin,G.M.
TITLE Sequencing of Drosophila chromosome 3R, region 99A7-99B5
JOURNAL Unpublished
2 (bases 1 to 311043)
REFERENCE
AUTHORS Celniker,S.E., Agbayani,A., Arcaina,T.T., Baxter,E., Blazej,R.G.,
Butenhoff,C., Champe,M., Chavez,C., Chew,M., Ciesiolka,L.,
Doyle,C.M., Farfan,D.E., Galle,R., George,R.A., Harris,N.L.,
Hinkle,A., Hoskins,R.A., Houston,K.A., Hummasti,S.R., Karra,K.,
Kearney,L., Kim,E., Lee,B., Lewis,S., Li,P., Ling,H.,
Moshrefi,A.R., Moshrefi,M., Nixon,K., Pacleb,J.M., Park,S.,
Pfeiffer,B., Poon,L., Richards,S., Sequeira,A., Sethi,H.,
Svirskas,R.R., Wan,K.H., Webster,D., Woolley,P., Yang,S., Yee,M.,
Zieran,L.L. and Rubin,G.M.
TITLE Direct Submission
JOURNAL Submitted (04-SEP-1999) Berkeley Drosophila Genome Project, MS
64-121, Lawrence Berkeley National Laboratory, One Cyclotron Road,
Berkeley, CA 94720, US
COMMENT On Sep 4, 1999 this sequence version replaced gi:5713324
gi:5670685.
Sequence submitted by:
Berkeley Drosophila Genome Project
Lawrence Berkeley National Laboratory, MS 64-121
Berkeley, CA 94720
For further information about this sequence, including its location
and relationship to other sequences, please visit our sequence
archive Web site (http://www.fruitfly.org/sequence/) or send email
to bdgp@fruitfly.berkeley.edu.
FEATURES
source
1..311043
/organism="Drosophila melanogaster"
/strain="y; cn bw sp"
/db_xref="taxon:7227"
/chromosome="3R"
/map="99A7-99B5"
/clone="BACS BACR48K01 (D488) and BACR48023 (D498)"
/clone_lib="RPCI-98 (Roswell Park Cancer Institute
Drosophila melanogaster BAC library, partial EcoRI in
pBACe3.6)"
/notes="BACR48K01 (D488) and BACR48023 (D498) were
sequenced as a project. BACR48K01 extends from bp 1 to bp
183,934. BACR48023 extends from bp 142,325 to bp
311,043."
BASE COUNT 87124 a 58661 c 68467 g 86791 t
ORIGIN
alignment_scores:
Quality: 743.00 Length: 154
Ratio: 4.921 Gaps: 0
Percent Similarity: 98.052 Percent Identity: 83.766
alignment_block:
US-09-525-867-1 x AC009888 ..
Align seg 1/1 to: AC009888 from: 1 to: 311043
60 GlyGluTyrValValAlaLysLeuAspLeuValasTrpAlaAArgAr 76
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
231343 GCGAGTGACCTGTGCCAGTGTGATGACCTGCTAAACCTGGGACGCAA 231392
76 gSerSerLeuTrpProMetThrPheGlyLeuAlaCysCysAlaValGluM 93
:::|||||:|||||:|||||:|||||:|||||:|||||:|||||:
231393 GGGATCCCTCGCGCGCTGACCTTTGTCTCGCTGCTGCGCGTGA 231442
93 etMetHisMetAlaAlaProArgTyrAspMetAspArgPheGlyValVal 109
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
231443 TGATGCACATTGGCGCTCTCGCTACGCATGGATCGCTACGGTGTGGTG 231492
110 PheArgAlaSerProArgGlnSerAspValMetileValalaglyThrLe 126
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
231493 TTCCGAGCATCTCTCGCGCGGAGCTGCTCATCTGCGCGGAAACCT 231542
126 tThrAsnLysMetAlaProAlaLeuArgLysValTyrAspGlnMetProG 143
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
231543 GACCAACAAGATGCGCGGCTTTCGGAAGATCTACGACCATGGCCG 231592
143 luProArgTyrValValSerMetGlySerCysAlaAsnGlyGlyTyr 159
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
231593 AGCCGAGATGGGTTCATTTCCATGGGCAGTTGCGCAATGGTGGCGGTAC 231642
160 TyrHisTyrSerTyrSerValValArgGlyCysAspArgIleValProVa 176
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
231643 TACCACTACTCTACTCGTGGTTCGCGGCTGCGATCGCATTTTCGGGT 231692
176 lasPileTyrIleProGlyCysProThrAlaGluAlaLeuLeuTyrG 193
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
231693 GGACATCTACGTGCGCGATGTCGCCGCCACCGCGAGGCGCTTAATGTACG 231742
193 lylleuGluLeuGlnArgLysIleLysArgGluArgArgLeuGlnIle 209
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
231743 GAATCCTCGAGTGCAGGAAGAAGGTAAAGCGCATGAGGACCGCTGCAGATG 231792
210 TrpTyrArgArg 213
|||||:|||||:
231793 TGGTACAGGAAG 231804
seq_name: gb.pl:BONADHPST
seq_documentation_block:
LOCUS BONADHPST 911 bp mRNA PLN 28-SEP-1995
DEFINITION B.oleracea mRNA for PSST subunit of NADH:ubiquinone oxidoreductase.
ACCESSION X82274
VERSION X82274.1 GI:562281
KEYWORDS NADH:ubiquinone oxidoreductase; NADH:ubiquinone oxidoreductase
subunit PSST.
SOURCE
ORGANISM Brassica oleracea.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Brassicales; Brassicaceae; Brassica.
REFERENCE
1 (bases 1 to 911)
Pogson,B.J., Downs,C.G., Davies,K.M., Morris,S.C. and
Buchanan-Wollaston,V.
Nucleotide sequence of a cDNA clone from broccoli with high
identity with the PSST subunit of NADH:ubiquinone oxidoreductase
Plant Physiol. 108 (2), 859-860 (1995)
95334517
MEDLINE
REFERENCE
2 (bases 1 to 911)
Pogson,B.J.
Direct Submission
TITLE Submitted (07-SEP-1994) B.J. Pogson, University of Arizona, Dept of
Plant Sciences, 303 Forbes Hall, Tuscon 85721, Arizona, USA
JOURNAL Location/Qualifiers
FEATURES
source
1..911
/organism="Brassica oleracea"
/cultivar="Shogun"
/db_xref="taxon:3712"
/dev_stage="mature plant"
/tissue_type="florets"
/clone="pBROC4"
/notes="L."
35..682
/codon_start=1

```

```

/product="PSST subunit of NADH: ubiquinone oxidoreductase"
/protein_id="CAA57725.1"
/db_xref="GI:562282"
/db_xref="SWISS-PROT:P42027"
/translation="MAMITNTATRLPLVLQSHRAAAVSHLHTSLPALSPAITPTSYT
RPGPSTSAIPPGSLKTAETVISKVDLMMNWRGSIWPMITGLACCAVEMHHTGAAR
YDLRFGILIRPSPQSDCMIVAGTLFNKAPALRKVDQMPPEPRWVISMGCANGGY
YYHYSYSVRGCDRIVPVDIYVPGCPPTAEALLYGLLQKKINRRKDLHWNK"
BASE COUNT      226 a      234 c      218 g      233 t
ORIGIN

alignment_scores:
  Quality: 736.00      Length: 181
  Ratio: 4.543        Gaps: 0
  Percent Similarity: 89.503      Percent Identity: 73.481

alignment_block:
US-09-525-867-1 x BONADHPST ..
Align seg 1/1 to: BONADHPST from: 1 to: 911
33  SerValAlaThrAspGlyProSerSerThrGlnProAlaLeuProLysAl 49
||| ||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
137 TCCTCCCGGACACACCGAGCTTCATACACAGCAGGTCCTCCATCGAC 186
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
49  aArgAlaValAlaProLysProSerSerArgGlyGluTyrValValAlaL 66
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
187 CTCGGCTCCTCCCGGGTCTCTCGAGAGCGGGAGTTCGTGATCTCAA 236
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
66  yLeuAspAspLeuValAsnTrpAlaArgArgSerSerLeuTrpProMet 82
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
237 AGGTAGATGATCTGATGAAGTACGCGCCGTAGGAGTATGATCTGCGCCATG 286
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
83  ThrPheGlyLeuAlaCysCysAlaValAlaGluMetMetHisMetAlaAlaPr 99
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
287 ACCITCGGGTCCTCGCTCGCGCGTGGAGATGATCATACAGGTGCGCG 336
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
99  oArgTyrAspMetAspArgPheGlyValValPheArgAlaSerProArgG 116
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
337 TCGCTACGATCTGGACCGGATTCGGTATCATCTTCAGGCGGAGTCTCGCC 386
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
116 lnSerAspValMetIleValAlaGlyThrLeuThrAsnLysMetAlaPro 132
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
387 AATCGGATGATGATGTCGCGGACGCTTACTAACAGATGCTGCC 436
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
133 AlaLeuArgLysValTyrAspGlnMetProGluProArgTyrValValSe 149
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
437 GCCTTCGCAAGTTTATGACAGATCCCGGACCCAGGTGGGTGATTTC 486
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
149 rMetGlySerCysAlaAsnGlyGlyTyrTyrHisTyrSerTyrServ 166
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
487 AATGGGAAGTTGTCCCAACGAGGTGGATATTATCATCTACTCTACTCGG 536
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
166 alValArgGlyCysAspArgIleValProValAspIleTyrIleProGly 182
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
537 TCGTTCAGGGTGGACAGAAATGTGCCAGTGGACATCATGCTCCCTGT 586
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
183 CysProThrAlaGluAlaLeuLeuTyrGlyIleLeuGlnLeuGlnAr 199
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
587 TGCCCTCCAAACAGCAGAGGCTCTGCTATGTGACTCTCCAGCTTCAGAA 636
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
199 gLysIleIleArgGluArgGluLeuGlnIleTyrTyrArgArg 213
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
637 GAAATCAACAGCGCAAGGATTCTTGCATCTGCTGGGAACAAG 679
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
seq_name: gb_pl.STPSST
seq documentation_block:
LOCUS STPSST 935 bp mRNA PLN 16-MAR-1999
DEFINITION S.tuberosum mRNA for NADH-ubiquinone oxidoreductase subunit.
ACCESSION X96671
VERSION X96671.1 GI:1235606

```

```

KEYWORDS      NADH-ubiquinone oxidoreductase; PSST gene.
SOURCE        Potato.
ORGANISM      Solanum tuberosum
REFERENCE     1 (bases 1 to 935)
AUTHORS      Heiser,V., Brennicke,A. and Grolmann,L.
TITLE         The plant mitochondrial 22 kDa (PSST) subunit of respiratory chain
              complex I is encoded by a nuclear gene with enhanced transcript
              levels in flowers
JOURNAL       Plant Mol. Biol. 31 (6), 1195-1204 (1996)
MEDLINE       97071689
REFERENCE     2 (bases 1 to 935)
AUTHORS      Heiser,V.
TITLE         Direct Submission
JOURNAL       Submitted (15-MAR-1996) V. Heiser, Inst.f. Genbiologische Forschung
              Berlin, Ihnestr. 63, D-14195, Berlin, FRG
FEATURES      Location/Qualifiers
               1..935
                /organism="Solanum tuberosum"
                /cultivar="Desiree"
                /db_xref="taxon:4113"
                /clone_lib="lambda ZAP II"
                /tissue_type="green leaves"
                50..691
                 /gene="PSST"
                 50..691
                  /gene="PSST"
                  /EC_number="1.6.5.3"
                  /note="22kDa mitochondrial complex I"
                  /codon_start=1
                  /product="NADH-ubiquinone oxidoreductase"
                  /protein_id="CAA65451.1"
                  /db_xref="GI:1235607"
                  /db_xref="SWISS-PROT:Q43844"
                  /translation="MALIARNKLDTAFLOPAAHTIHTLPSLQOPASSPATSGG
                  AQPSMNTFAGISPAEYISKVDLMMNWRGSIWPMITGLACCAVEMHHTGAARID
                  FDRFGILIRPSPQSDCMIVAGTLFNKAPALRKVDQMPPEPRWVISMGCANGGY
                  HYSYAVRGCDRIVPVDIYVPGCPPTAEALLYGLLQKKINRRKDLMMWTQ"
                  50..142
                   /gene="PSST"
                   /evidence=experimental
                   143..588
                    /gene="PSST"
                    583..888
                     BASE COUNT      221 a      212 c      221 g      281 t
                     ORIGIN

alignment_scores:
  Quality: 732.50      Length: 206
  Ratio: 4.234        Gaps: 4
  Percent Similarity: 83.981      Percent Identity: 67.961

alignment_block:
US-09-525-867-1 x STPSST ..
Align seg 1/1 to: STPSST from: 1 to: 935
17  LeuArgSerSerValGlyProAlaValGln...AlaArgGlyValHisGl 32
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
77  CTGCTCACCGAAGCGCGCGTCTCTCAAGGCGCGCGACGATCCACAC 126
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
32  nSerValAlaThrAspGlyProSer...SerThrGlnProAlaLeuProL 48
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
127 GACACTC.....CCATCTCTCTCACAGCAACCTGCTCTCTC 164
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
48  ysAlaArgAlaValAlaProLysPro..... 56
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
165 CGGCCACTTCAGGTGGCGACACGCTCCGTCCTCCATGATACACGAGGG 214
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
57  ...SerSerArgGlyGluTyrValValAlaLysLeuAspLeuValAs 72

```



```

:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
601 CTCCTCCAATTCAGAGAGAGATCAATCGGCGCAAGGACITCCTCCTGTG 650
210 PTYRARGATG 213
|:::|
651 GTGGACAAG 660
```

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 12, 2002, 17:41:45 ; Search time 1339.84 seconds  
(without alignments)  
10145.746 Million cell updates/sec

Title: US-09-525-867-9

Perfect score: 824

Sequence: 1 cggctgagcggctcgagcg.....cgggcaaaaaaaaaaaaaa 824

Scoring table: OLIGO\_NUC

Gapop 60.0 , Gapext 60.0

Searched: 1472140 seqs, 8248589755 residues

Word size : 0

Total number of hits satisfying chosen parameters: 2944280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : GenEmbl.\*

1: gb.ba.\*  
2: gb.htg.\*  
3: gb.in.\*  
4: gb.om.\*  
5: gb.ov.\*  
6: gb.pat.\*  
7: gb.ph.\*  
8: gb.pl.\*  
9: gb.pr.\*  
10: gb.ro.\*  
11: gb.sts.\*  
12: gb.sy.\*  
13: gb.un.\*  
14: gb.vi.\*  
15: em.ba.\*  
16: em.fun.\*  
17: em.hum.\*  
18: em.in.\*  
19: em.om.\*  
20: em.or.\*  
21: em.ov.\*  
22: em.pat.\*  
23: em.ph.\*  
24: em.pl.\*  
25: em.ro.\*  
26: em.sts.\*  
27: em.sy.\*  
28: em.un.\*  
29: em.vi.\*  
30: em.htgo\_hum.\*  
31: em.htgo\_inv.\*  
32: em.htgo\_rod.\*  
33: em.htg\_hum.\*  
34: em.htg\_inv.\*  
35: em.htg\_rod.\*  
36: em.htg\_other.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

8

#### REMARK

NIH-MGC Project URL: <http://mgc.nci.nih.gov>

Result No.	Score	Query Match	Length	DB	ID	Description
1	736	89.3	794	9	BC001715	BC001715 Homo sapi
2	700	85.0	787	9	BC005954	BC005954 Homo sapi
3	366	44.4	754	9	AF060512	AF060512 Homo sapi
4	190	23.1	34875	9	AC005329	AC005329 Homo sapi
5	38	4.6	752	4	BTPSSTSU	X65020 B.taurus m
6	38	4.6	329709	1	AP002997	AP002997 Mesorhizo
7	35	4.2	226013	2	AC073680	AC073680 Mus muscu
8	28	3.4	68808	2	AC022708	AC022708 Homo sapi
9	28	3.4	172631	2	AC012566	AC012566 Homo sapi
10	28	3.4	198491	2	AC022249	AC022249 Homo sapi
11	27	3.3	789	6	AX106251	AX106251 Sequence
12	27	3.3	789	6	AX140542	AX140542 Sequence
13	27	3.3	1038	6	AX008551	AX008551 Sequence
14	27	3.3	1038	9	HOMLBP	L48925 Homo sapien
15	27	3.3	1879	10	AF068261	AF068261 Rattus no
16	27	3.3	3551	10	D88190	D88190 Rattus norv
17	27	3.3	158983	2	AC020598	AC020598 Homo sapi
18	27	3.3	179206	9	CNS01DS6	AL121656 BAC seque
19	27	3.3	200368	9	CNS01DUH	AL133244 BAC seque
20	26	3.2	897	8	AF281035	AF281035 Lupinus l
21	26	3.2	911	8	BONADHPST	X82274 B.cleracea
22	26	3.2	3739	1	PDENOUVR	M93015 Paracoccus
23	26	3.2	13614	1	AF029365	AF029365 Rhodobact
24	26	3.2	138705	2	AC068923	AC068923 Oryza sat
25	26	3.2	144201	2	AC091774	AC091774 Oryza sat
26	26	3.2	153673	2	AP003769	AP003769 Oryza sat
27	25	3.0	2924	8	AF279902	AF279902 Lupinus l
28	25	3.0	144593	8	AC021891	AC021891 Genomic S
29	25	3.0	164192	9	AC074338	AC074338 Human Chr
30	25	3.0	179488	9	AC083863	AC083863 Homo sapi
31	24	2.9	224	8	AY022669	AY022669 Oryza sat
32	24	2.9	224	8	AY022809	AY022809 Oryza sat
33	24	2.9	798	9	HSA250562	AJ250562 Homo sapi
34	24	2.9	1792	9	HUMMRNAB	L10373 Human (clon
35	24	2.9	1829	9	AB062057	AB062057 Homo sapi
36	24	2.9	6043	9	AB037720	AB037720 Homo sapi
37	24	2.9	10311	1	AB008339	AB008339 Agrobacte
38	24	2.9	10898	1	AP005781	AP005781 Caulobact
39	24	2.9	116275	2	AP003993	AP003993 Oryza sat
40	24	2.9	122215	2	AP003811	AP003811 Oryza sat
41	24	2.9	139487	2	AP003513	AP003513 Oryza sat
42	24	2.9	141017	8	AC037197	AC037197 Oryza sat
43	24	2.9	142101	2	AC091811	AC091811 Oryza sat
44	24	2.9	144798	8	AC026758	AC026758 Oryza sat
45	24	2.9	178024	8	AC025783	AC025783 Oryza sat

#### ALIGNMENTS

RESULT 1  
BC001715  
LOCUS BC001715 794 bp mRNA PRI 12-JUL-2001  
DEFINITION Homo sapiens, Similar to CG9172 gene product, clone MGC:886  
IMAGE:3503064, mRNA, complete cds.  
ACCESSION BC001715  
VERSION BC001715.1 GI:12804588  
KEYWORDS MGC.  
SOURCE human.  
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 794)  
Strausberg, R.  
Direct Submission

REFERENCE  
AUTHORS  
TITLE  
JOURNAL

Submitted (16-JAN-2001) National Institutes of Health, Mammalian  
Gene Collection (MGC), Cancer Genomics Office, National Cancer  
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,  
USA

## COMMENT

Contact: MGC help desk  
Email: cgabbs-r@mail.nih.gov  
Tissue Procurement: ATCC  
cDNA Library Preparation: Rubin Laboratory  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Genome Sequence Centre,  
BC Cancer Agency, Vancouver, BC, Canada  
info@bcgsc.bc.ca  
Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield,  
Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin,  
Leticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo  
Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven  
Ness, Pawan Pandoh, Anna-Liisa Prabhu, Parvaneh Saedi, Jacqueline  
Schein, Duane Smallus, Michael Smith, Lorraine Spence, Jeff Stott,  
Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy,  
George Yang, Scott Zuyderduyn, Marco Marra.

Clone distribution: MGC clone distribution information can be found  
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
Series: IRAL Plate: 6 Row: b Column: 12  
This clone was selected for full length sequencing because it  
passed the following selection criteria: matched mRNA gi: 7688970.

## FEATURES

## source

1. .794  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="MGC:886 IMAGE:3503064"  
/tissue\_type="Brain, neuroblastoma"  
/clone\_lib="NIH\_MGC\_19"  
/lab\_host="DH10B-R"  
/note="Vector: pORF7"  
6. .647  
/codon\_start=1  
/product="Similar to CG9172 gene product"  
/protein\_id="AA01715.1"  
/db\_xref="GI:12804589"  
/translation="MAYISAPGLRGLRSLSSVGPVAVQGVHQSVATDGPSSSTP  
ALPKAVAPKSPSSGEYVAKLDLVNWRSSLWPTFLACCAVMHMAAPRYD  
MORFYGVPASRQSDMIVAGLTNKPALRYVDOMPEPVYVWSGSCANGGY  
HYSYVRCSDRIVPDVIFGCPPTABALLYILQIKRRLRIQIYWR"  
BASE COUNT 170 a 262 c 236 g 126 t  
ORIGIN

## Query Match

Best Local Similarity 89.3%; Score 736; DB 9; Length 794;

Matches 736; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 74 ccaagatggcgggtgctgagctcctggcctgcggcgtccgagatccttgctgcgct 133  
Db 1 CCAGATGGCGGTGCTGTACGTCCTCCGCTGCGCGCTCCGAGTCTGTGCTGCGCT 60  
Qy 134 ccagctgggcccggctgtgagcagcagaggtgtccatcagagcgtggccaccgagtc 193  
Db 61 CCAGCGTGGGCGCGGTGTGCAGGCACGAGGTGTCCATCAGAGCGTGGCCACCGATGCC 120  
Qy 194 caagcagaccagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 253  
Db 121 CAACACACCCAGCCAGCCCTCCAAAGCCAGAGCCGTGCTCCAAACCCAGCAGCC 180  
Qy 254 gggcgagtagtggtggcgaagctggatgacctctcaactggcggcggcggagattctc 313  
Db 181 GGCGCAGATATGTGGTGGCCAAAGCTGGATGACCTGCTCAACTGGCGCGCGGAGTTCTC 240  
Qy 314 tgtggccatgacctggcctgagcagcagcagcagcagcagcagcagcagcagcagc 373  
Db 241 TGTGGCCCATGACCTTGGCGCTGGCTGTGGCGGTGGAGATGATGACATGGCAGCAC 300  
Qy 374 ccgctacgacatggagccgcttggcggtgtgtcttcgcgcagccgcgcagtcgcgagc 433  
Db 301 CCCGCTACGACATGGACCGGCTTTGGGTGTCTTCGCGCCAGCCGCGCCAGTCCGAGC 360  
Qy 434 tcatgatcgtggcggcagcactcaccacaaagatggccccagcgttcggaagtcctacg 493

Db 361 TCATGATGTCGGCCGCACACTCACCACAAGATGCCCCAGCGTTTCGCAAGGTCTACG 420  
Qy 494 accagatggcggagcgcgctacgtgtgtctccatggggagctgcgcacagagagagct 553  
Db 421 ACCAGATGCGGAGCGCGCTACGTGTCTCCATGGGAGCTGCACCAACGAGAGGCT 480  
Qy 554 actaccactattctactcgggtggtgaggggtgagcgcagcagcagcagcagcagc 613  
Db 481 ACTACCCTATTCTTACTCGGTGTGAGGGGTGCGACCGCATCGTCCCGTGGACATCT 540  
Qy 614 acatccagagctgccacactacgcgcgagcgcctgctctacgcgcatcctgcagctgcaga 673  
Db 541 ACATCCAGGCTGCCACCTACGCGCGAGCGCCGTGCTCTACGGCATCTCTGACAGTGCAGA 600  
Qy 674 ggaagatcaaggggagcggagcgtcagatctgttaccgcgggtgagcgcgcgcgcgc 733  
Db 601 GGAAGATCAAGCGGAGCGAGGCTGCAGATCTGTGTACCGAGGTAGCGCCGCCGCC 660  
Qy 734 gcgcgcgagcctgtgcgcgtctctccccagcctgtctcccgtagagtttcaat 793  
Db 661 CGCGCGGAGCCCTGTGCGGTCTCTGTCGCCAGCCTGTGTGTCGGTGAGGTTGTCAAT 720  
Qy 794 aaactcgccctcgggc 809  
Db 721 AAACCTGCCCTCGGC 736

## RESULT 2

## BC005954

## LOCUS

## DEFINITION

## ACCESSION

## VERSION

## KEYWORDS

## SOURCE

## ORGANISM

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## REMARK

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

CDS	/note="vector: pDNR-LIB"
	21...662
	/codon_start=1
	/product="Unknown {protein for MGC:14592}"
	/protein_id="AAH05954.1"
	/db_xref="GI:13543603"
	/translation="MAVLASGLRFRILGRSSVGLAVQARGVHQSVATDGSPSSTOP ALPKARAVAPKPSRGVEYVAKLDLIVNWRSSLPMTFLGACCAEMMHMAAPRD MDRFVVFRASPQSDEVMIVAGTLINKMAPLKRVDDPFRFVWSMGSCANGGGY HYGSYVGRCDRIYVDIITPCPGTEAFALLYILLQORKIKRRRLQIWRR"
BASE COUNT	152 a    262 c    244 g    129 t
ORIGIN	
	Query Match                 85.0%; Score 700; DB 9; Length 787;
	Best Local Similarity      99.9%; Pred. No. 0;
	Matches 750; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY	59 gtctaaagccgagcgaagtggcggtgtgttcagctctgcctggcgtgcggcttcccga 118
Db	1 GTCTAAGGCCAGGCCAAGATGCGGTGTGTTCAGCTCCTGGCCTGC GGCTTCCGGA 60
QY	119 tccttgctcgctccagctgggccggcgtgtgaggcaagaggtgtccatcacagcg 178
Db	61 TCCTTGCTGGCTCCAGCTGGCGCTGGCTGTGCAGSCACGAGGTGTCCATCAGACG 120
QY	179 ttggcacagatggccccagcagaccacgcctgcctgtccaaagcgccagagccgtggctc 238
Db	121 TGCCACCAGATGGCCCAAGCAGCACCCAGAGCTGCCCTGCCAAGGCCAGAGCGGTGGGTC 180
QY	239 ccataccagcagcggggcgagtattgtgtgccaaagctgatgacctcgtccaactggg 298
Db	181 CCAAACCCAGCAGCGGGCGGCGAGTATTGTGGTGCCAAAGTGGTAGCTCTCAACTGGG 240
QY	299 ccgcgcgagttctctgtggccaatgaccttcggcctgtgcctgtgcgcgtggagatga 358
Db	241 CCGCGCGAGTTCTCTGTGGCCCCATGACCTTCGGCCTGGCCTGCTGGCGCGTGGAGATGA 300
QY	359 tgcatatggcagcaccgccttacqacatgacacgctttggcgtgtctccgcgcacgcc 418
Db	301 TGCACATGGCAGCACCCCGCTACGACATGACACGCTTGGCGTGTCTTCGCGCCAGCC 360
QY	419 cgccgcagtcgcagctcatatgctdggccggcacactcaccacaagaatggtcccacgcgc 478
Db	361 CGCGCCAGTCCGCAGCTCATGATCGTGGCCGGCACACTCCACAAAGATGGCCCGCAGGC 420
QY	479 ttgcgaagctctacgaccagatgcggagcgcgcgtacgttgtctccatgggagctcgc 538
Db	421 TTCGAAGGTCTACGCCAGATGCGGAGGCGCGGTACATGTGTCTCCATGGGAGCTCGG 480
QY	539 caaacggaggagctactaccactattcctaactcgttgttgaggggctgcgaccgcatcg 598
Db	481 CCAAAGGAGGAGGCTACTTACCACATTCTCTACTCGGTGTGGGGGCTCGCAGCCGATCG 540
QY	599 tgcggtggacatctacatccaggctgcaccaactacggcagagggcctgtctcaggca 658
Db	541 TGCCCGTGGACATCTACATCCCAGCTTGCCACTACGGCCGAGGCGCTGTCTACGGCA 600
QY	659 tcctgcagctgcagaggagaatatcaagcggagcggaggtgcagatctcgttaccgcaggt 718
Db	601 TCTCGAGCTGCAGAGGAAGATCAAGCGGAGGCGGAGGCTGCAGATCTGGTACCGCAGGT 660
QY	719 agc 778
Db	661 AGCGCGCGCGCGCGCGCGCGAGCCTGTGCGCGTCTGTCCCAAGCAGCTGTGTGTCC 720
QY	779 cgtgaggttgttaataaaacctgcctcgggc 809
Db	721 CGTGAGGTTGTCAATAAACCTGCCTCGGGC 751

RESULT  
AF060512  
3

LOCUS	AF060512	754 bp	mRNA	PRI	02-JAN-2001
DEFINITION	Homo sapiens clone 016d03 My017 protein mRNA, complete cds.				
ACCESSION	AF060512				
VERSION	AF060512.1 GI:12001973				
KEYWORDS	.				
SOURCE	human.				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
AUTHORS	Mao,Y.M., Xie,Y., Huang,X.Y., Ying,K. and Dai,J.L.				
TITLE	Direct Submission				
JOURNAL	Submitted (20-APR-1998) Institute of Genetics, School of Life Science, Fudan University, 220 Handan Rd., Shanghai 200433, P.R.China				
FEATURES	Location/Qualifiers				
Source	1..754				
	/organism="Homo sapiens"				
	/db_xref="taxon:9606"				
	/clone="016d03"				
	/tissue_type="brain"				
	/dev_stage="fetus"				
	10..654				
CDS	/codon_start=1				
	/product="My017 protein"				
	/protein_id="AA043130.1"				
	/db_xref="GI:12001974"				
	/translation="MAVLSAPGLRGRFLGRSSVGLVQARGVHQSVATDGSSTSPD ALPKRAVAPKPSRGEYVAKLDLVNWRKSSLPWTFGLACCAVEMHMAAPRID MDREGVYRASPQSDVMVTAQTLNKPALRKSPPPDAGALRLHGLRQRLRGL LPFLGGLGELRPHRARGHLHPRLPTTAEALLYGLQLQRKIKRRRLQIWRP"				
BASE COUNT	120 a	267 c	239 g	128 t	
ORIGIN					
Query Match	44.4%; Score 366; DB 9; Length 754;				
Best Local Similarity	95.8%; Pred. No. 2e-171;				
Matches	416; Conservative	0; Mismatches	1; Indels	0; Gaps	0;
QY	70	gagggcaagatgagcgtgctgtagctccctggcctgcgagcgttcgagatccttgctg	129		
Db	1	GAGGCCAAGATGCGCGTGTCTGACATCCTGGCCTTCGGCGGCTTCGGATCCTTGGTCTG	60		
QY	130	cgctcagcgtgggccccggctgtgcaggcacgaggtgtccatcagagcgtggccaccgat	189		
Db	61	CGCTCCAGCGTGGCGCTGGCTGTGCAGGCACAGAGTGTCCATCAGACGCTGGCCACCGAT	120		
QY	190	ggcccaagcacaccacgcctgcctcccaaaagccagacgctggtctcccaaacccagc	249		
Db	121	GGCCCCAAGCAGCAGCCACGCCCTGCCCTGCCAAAGCCAGAGCGGTGCTCCCAACCCAGC	180		
QY	250	agccggggcgagatgtggtggcgaagctggatgacctgctcaactggccgcgcgcggagt	309		
Db	181	AGCCGGGGGAGTATGTGTGTGGCCAAAGCTGGATGACCTGCTCAACTGGCGCGCGGAGT	240		
QY	310	tctctgtgccaatgaccttgagccttgagcctgtgcgcgtggagatgatgcacatggca	369		
Db	241	TCTCTGTGCCCATGACCTTCGGCCTGGCTGTGCGCTGGAGATGATGCACATGGCA	300		
QY	370	gcaccgcgtacgacatggaccgctttggcgtggtcttcgcgcagccgcgcgcagctcc	429		
Db	301	GCACCCCGTAGCATGACATGGACCGCTTTGGCGTGGTCTTCGCGCGCAGCCGCGCCAGTCC	360		
QY	430	gacgctgatcgtggccgggcacactcaccacaagaatggccccagcgttcgcgaag	486		
Db	361	GACGTATGATGTGGCGGCACACTCACCACAAGATGGCCCCAGCGCTTCGCCAAG	417		
RESULT	4				
AC005329	AC005329	34875 bp	DNA	PRI	28-JUL-1998
LOCUS	Homo sapiens chromosome 19, cosmid R34382, complete sequence.				
DEFINITION	Homo sapiens chromosome 19, cosmid R34382, complete sequence.				
ACCESSION	AC005329				

## RESULT 4

RESULT 3  
AF060512

AC005329	LOCUS	AC005329	34875 bp	DNA	PRI	28-JUL-1998
	DEFINITION	Homo sapiens chromosome 19, cosmid R34382, complete sequence.				
	ACCESSION	AC005329				







AP002997 BA000012  
 AP002997.2 GI:14022051  
 Mesorhizobium loti (strain:MAFF303099) DNA.  
 Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;  
 Phyllobacteriaceae; Mesorhizobium.  
 1 (sites)  
 Kaneko,T., Nakamura,Y., Sato,S., Asamizu,E., Kato,T., Sasamoto,S.,  
 Watanabe,A., Iidesawa,K., Ishikawa,A., Kawashima,K., Kimura,T.,  
 Kishida,Y., Kiyokawa,C., Kohara,M., Matsumoto,M., Matsuno,A.,  
 Mochizuki,Y., Nakayama,S., Nakazaki,N., Shimpo,S., Sugimoto,M.,  
 Takeuchi,C., Yamada,M. and Tabata,S.  
 Complete genome structure of the nitrogen-fixing symbiotic  
 bacterium Mesorhizobium loti  
 DNA Res. 7 (6), 331-338 (2000)  
 21082930  
 2 (bases 1 to 329709)  
 Kaneko,T.  
 Direct Submission  
 Submitted (05-DEC-2000) Takakazu Kaneko, Kazusa DNA Research  
 Institute, The First Laboratory for Plant Gene Research; Yana  
 1532-3, Kisarazu, Chiba 292-0812, Japan  
 (E-mail:kaneko@kazusa.or.jp/rhizobase/)  
 URL: <http://www.kazusa.or.jp/rhizobase/>  
 Tel:81-438-52-3935(ex.2338), Fax:81-438-52-3934)  
 On May 11, 2001 this sequence version replaced gi:11994965.  
 Location/Qualifiers  
 1..329709  
 /organism="Mesorhizobium loti"  
 /strain="MAFF303099"  
 /db\_xref="taxon:381"  
 complement(79..402)  
 /gene="mlr1244"  
 complement(79..402)  
 /note="unknown protein"  
 /codon\_start=1  
 /transl\_table=11  
 /protein\_id="BAB48662.1"  
 /db\_xref="GI:14022052"  
 /translation="MSRKGSLIPFGNSFHYWADCLMVQRPIMSDLLLSIFTFTVMR  
 LLKGWLRNPQYLATILGAVVAMLLNGLWPAYDDDFVIGGVGTGFGWAGNALFEDA  
 ILGIA"  
 411..2480  
 /gene="mlr1245"  
 411..2480  
 /gene="mlr1245"  
 /codon\_start=1  
 /transl\_table=11  
 /product="probable electron transfer protein"  
 /protein\_id="BAB48663.1"  
 /db\_xref="GI:14022053"  
 /translation="MNSPANTIDPLVFMPSKGRGPRVGTPTVLDARQLGVYVESVC  
 GGRVCGQCEVEGNGFAKHKIVSSNDHISPGPKKEERYVRGPERRLSCSAQI  
 LGDLVDPQVETINAOITRKDADTVIARDTVAIRMCYVEIEEPDMHKPLGLDRLKI  
 YMLADWGFKNLEFFYLLPOVQILIRKGNWTATAAIHKDADSDIARIYALPGLKNEA  
 YLACDIGSTIAHLYSLGSRVAASGTSNPOIRGEDLMKRSVYVMNPDGREGM  
 TVAVREATSSIVKCAEGNVQRNDILDSVFNPIWHHLFLGIDPTELGGAPFALAV  
 SGAVRIKASDGLKINOCARLYMLPCITAGHVGADAAVTISEGPHQDEMLLYDGT  
 NAEIYLNRARVANSPTGAPFGAIEISGQRAAGAIERKIDPTELPEPKYRIGS  
 ELWSDPGLVSGATGVTGICGSGIIEIYAENYLAGITISEDGVGDSLAARSPRVTA  
 NGRFYSYKGECEKITITQDVRAIQAKAALYAGITKLMKEQNTDHDRIHAFAGF  
 GSFIDPKYAMVGLITPCDDLDKSAVGNAGAGAMALLNRGYRRIETETVSQIEKIE  
 TALEPKFQEHFYAMALPNKVDPPFKLSAAVKLPKPKTVSEDDGIAGDTPRRSREGH  
 AARRSRE"  
 2529..2852  
 /gene="mlr1246"  
 2529..2852  
 /gene="mlr1246"  
 /note="unknown protein"  
 /codon\_start=1  
 /transl\_table=11

/protein\_id="BAB48664.1"  
 /db\_xref="GI:14022054"  
 /translation="MFPDRTDPTETTFLSRNLSDTQWHLPRIERPERPEILRRSSRE  
 MKMGNFNLGRFGVTIRQHARNKAIKALNSLFAHVQKIDGWSPSPDPQAALASLL  
 LGTAR"  
 3087..4040  
 /gene="mlr1247"  
 3087..4040  
 /gene="mlr1247"  
 /codon\_start=1  
 /transl\_table=11  
 /product="esterase"  
 /protein\_id="BAB48665.1"  
 /db\_xref="GI:14022055"  
 /translation="MPSIKSHFVSFLYSRKFAFSSPENLRBWLIAARKTEDHHPA  
 SLHORLHIQRTVDGFFYETIAPKAGQRRILHGAAYVFEITPPHWHLIAMADRLL  
 GIGTPIYPIAPEDHDFHAFMGVGDVIRQMLDETAEDIVFMGDDGAGNMVAVLTMM  
 AAEEGLPLPAHLYISPLDMSLNPKLFAERNRPNWLGAGLEAIRMYSGIDRSR  
 WHISPIYGDVSLPTLLTGTSHDLTLIFAQKARDAGVEVEVVEEGMFWHAPL  
 IDMPERRARQHIIVAFILGENRSPVRQDGKTRFOAPSAAAE"  
 complement(4062..4961)  
 /gene="mlr1248"  
 complement(4062..4961)  
 /gene="mlr1248"  
 /codon\_start=1  
 /transl\_table=11  
 /product="probable transcriptional regulator"  
 /protein\_id="BAB48666.1"  
 /db\_xref="GI:14022056"  
 /translation="MNAPLNHPLLDLVLTFTVAETGSETTAAAVERTPSPAVS  
 MOIKLEDILGSRVFPARDARSVSLTGTGEMLLGYARLLSINREVYSKFIPLVGV  
 RLGSFDDYGERVLPVILKRFQAQSHPSIADVTIDQSNLRRRMDRALDILLITNYSK  
 TSALGAEVLLTPEIVWAGKQCAHLREPLFVSLWEEGCARAGALEALRGRENYRV  
 AYSATAGAQORAAIMADLAVAPLPRSEFLGNDMVELCPKDGMPDITGTYNLAMVYAPAS  
 APKAVADHIRATEFEVFRGTGF"  
 5113..5379  
 /gene="msr1249"  
 5113..5379  
 /gene="msr1249"  
 /note="unknown protein"  
 /codon\_start=1  
 /transl\_table=11  
 /protein\_id="BAB48667.1"  
 /db\_xref="GI:14022057"  
 /translation="MGLKHAACKTSQIKRNVPGPAPKENEMSILSSGLRIATEFSA  
 RARYQTBRAIRSLPIELQKIDGWPEASDTGTGRNGVSGWAGAK"  
 6136..7137  
 /gene="mlr1251"  
 6136..7137  
 /gene="mlr1251"  
 /codon\_start=1  
 /transl\_table=11  
 /product="probable transcriptional regulator"  
 /protein\_id="BAB48668.1"  
 /db\_xref="GI:14022058"  
 /translation="MNAIKHPKIRSFVFFLVPDFMTAFATLDPRLSRANRLMGLEYA  
 RWRLSIDGKVRASNGECVANTSLPEERKKNAGDRPNMAIVCGSVNRYQNKSA  
 FAWLREENVRAVGGCTGAHTILAAALSNKRCAIHWENLPGFSEAPKANVPADL  
 FEVDQNYTCAGGTAALDMKLKIGDDDFDSLNVRYCEOVLTDRVSPDRORLPURA  
 RLGVQNSKYITIIELMEGNLSEPLSLIEIADHDVLSRRQIERLFRTEMGRSPARYLE  
 IRLDRAHLLIQSSMPVVEVAVACGVSASHFSKCIREFLYARSFQQRVDRKQLAA"  
 complement(7153..7956)  
 /gene="mlr1253"  
 complement(7153..7956)  
 /gene="mlr1253"  
 /codon\_start=1  
 /transl\_table=11  
 /product="ATP-binding protein of ABC transporter"  
 /protein\_id="BAB48669.1"  
 /db\_xref="GI:14022059"  
 /translation="MEPGVSKLIEGVSRITFAGVGRGLPVKALMPIDLVAANDFTITI  
 LGPSGCKSTLIRIVAGLEAPSEGRVLLDDKAVTRPGPDGRGMVFQSYTLFPMLTVSEN  
 IAFGLRGRMAPKERDDIVASYVDLVGLKGFENHWPQKLSGGMQQRTAIRALANDE

```

gene      ILLDPPGALDNCRTGLNQELLGIGWERRKXIVLVFVTHDIEAFMASRVIVMTARP
          GSINSDAIDLPFRPHYTKTSPFSALKARLTEDIRVEAMRTAQQA"
          complement(7941. .8720)
CDS       /gene="m111254"
          complement(7941. .8720)
          /gene="m111254"
          /codon_start=1
          /transl_table=11
          /product="permease protein of ABC transporter"
          /protein_id="BAB48670.1"
          /db_xref="GI:14022060"
          /translations="MKMRPLHPVSPGLRTGLGISFFVLFIAFWAMITLGGHVRIFLA
          DPLMDRLGRVLEDVDFWLDILITWRVGGFVLASVAVPIGAMGWKPEAFLE
          PFVSFARYLPASAFIPLLIILWAGIGLEKLIIVFGSVFOIILIVAVKVGSTRDLVE
          AAYTLSTNNIGYKRVIMANAPAEIATETRLVLGNWYIVVIAELIGSSGIGYMIIN
          SQRLARQGIIFGILVIGLIGLISDFAFRFRNLPFWSLA"
          complement(8802. .9752)
          /gene="m111256"
          complement(8802. .9752)
          /gene="m111256"
          /note="hypothetical protein"
          /codon_start=1
          /transl_table=11
          /protein_id="BAB48671.1"
          /db_xref="GI:14022061"
          /translations="MKRIAVFAATLTFAGLGAAPVMAATSVTIGISGWTGAPLTLAK
          QAGLEKGGIDVTLKKVPOASRLATASGDLOCATTVETVLVWNASGVTTKQIFOLD
          KSVGAGIYVRNDIKKIVADLKGNVASSAPGSPFELLAWLNKNGMSTKDYVFNLE
          PRAAQFLAGNDAAETPEFISAVRADOQHILATLDIPWVDIVGTGPEPLKA
          NPAAKALADSYEALDLKDKPKQSYEIMGADVOSAKEFEDSAYKYLKWADKANQQ
          FFTKEQDFSKTAGILLQMLGKIEAPDVATLADTSAVAN"
          complement(9963. .11000)
          /gene="m111257"
          complement(9963. .11000)
          /gene="m111257"
          /note="unknown protein"
          /codon_start=1
          /transl_table=11
          /protein_id="BAB48672.1"
          /db_xref="GI:14022062"
          /translations="MSSISARPIRSKPSVATSREMTRTSSRAADSGGPRDLLQKP
          GFEPGCLPQVVFAYPRGMSALSRAHLLSCGALALMLWPGSRADPEVDVLELV
          ADVYSLSMADELEIORHYAALHDNVLOIADGANGKIAVTVYEWAGTWQRVIV
          FWTIVNRADEARVAAQLGAPPNARRTISIGALEFGSDLFAESYIGTKRVIDISG
          DPNNGQAPVNTDGVTRGVQIVINGLPLMTGRGLSGAYDVNDLDRYSDCVIGPGA
          FMIPVNDWTQFPPIARRKLVLLELAGPASPQWAAEADHPVVLTDCKPAADCQIGEM
          WRNRNWLDSR"
          complement(10906. .13467)
          /gene="m111258"
          complement(10906. .13467)
          /gene="m111258"
          /codon_start=1

Query Match      4.6%; Score 38; DB 1; Length 329709;
Best Local Similarity 100.0%; Pred. No. 2.2e-08;
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 481 qcaaggtctacgacagatgcgagcgccgctactgt 518
      |||||
DB 102691 CCAGAGTCTACGACGAGATGCCGCGCGCTACGT 102654

RESULT 7
AC073680/c      AC073680 226013 bp DNA HTG 29-JUN-2000
LOCUS           Mus musculus clone RP23-116M1, WORKING DRAFT SEQUENCE, 20 unordered
DEFINITION      pieces.
ACCESSION      AC073680
VERSION        AC073680.1 GI:8810297
KEYWORDS       HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE         house mouse.
ORGANISM       Mus musculus
                Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
```

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 226013)  
DOE Joint Genome Institute.  
Sequencing of Mouse  
Unpublished  
2 (bases 1 to 226013)  
DOE Joint Genome Institute.  
Direct Submission  
Submitted (29-JUN-2000) Production Sequencing Facility, DOE Joint  
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA  
-----Genome Center  
Center: Joint Genome Institute  
Center Code: JGI  
Web site: <http://www.jgi.doe.gov>

Project Information  
Center Project Name: 1781705  
Center clone name: RPCI-23\_116M1  
-----

## Summary Statistics

Consensus quality: 209118 bases at least Q40  
Consensus quality: 215883 bases at least Q30  
Consensus quality: 217616 bases at least Q20  
Estimated insert size: 216000; agarose-fp estimation  
Estimated insert size: 224113; sum-of-contigs estimation  
Quality coverage: 9.65 in Q20 bases; agarose-fp estimation  
Quality coverage: 9.3 in Q20 bases; sum-of-contigs estimation  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 20 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.

1 1157: contig of 1157 bp in length  
1257: gap of unknown length  
1258: contig of 1051 bp in length  
1258: contig of 1051 bp in length  
2308: contig of 1051 bp in length  
2408: gap of unknown length  
3408: contig of 1000 bp in length  
3508: gap of unknown length  
3509: contig of 1257 bp in length  
4785: contig of 1257 bp in length  
4855: gap of unknown length  
6182: contig of 1317 bp in length  
6282: gap of unknown length  
7697: contig of 1415 bp in length  
7797: gap of unknown length  
9195: contig of 1398 bp in length  
9295: gap of unknown length  
10579: contig of 1284 bp in length  
10679: gap of unknown length  
10680: contig of 1394 bp in length  
12074: gap of unknown length  
12174: contig of 1511 bp in length  
13684: gap of unknown length  
13784: gap of unknown length  
14932: contig of 1148 bp in length  
15032: gap of unknown length  
17640: contig of 2608 bp in length  
17740: gap of unknown length  
19342: contig of 1602 bp in length  
19442: gap of unknown length  
21366: contig of 1924 bp in length  
21466: gap of unknown length  
21367: contig of 5780 bp in length  
27246: gap of unknown length  
27346: gap of unknown length  
30988: contig of 3642 bp in length  
31088: gap of unknown length  
35750: contig of 4662 bp in length  
35850: gap of unknown length  
35851: contig of 14839 bp in length  
50789: gap of unknown length  
50790: contig of 19485 bp in length  
70274: contig of 19485 bp in length  
70275: gap of unknown length

```

FEATURES
  source      * 70375 226013: contig of 155639 bp in length.
                Location/Qualifiers
                1..226013
                /organism="Mus musculus"
                /db_xref="taxon:10090"
                /clone_lib="RP23-116Mi"
                /clone_lib="RPCI mouse BAC library 23"
BASE COUNT    54098 a 60347 c 59323 g 50104 t 1941 others
ORIGIN

Query Match
Best Local Similarity 4.2%; Score 35; DB 2; Length 226013;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 51 cagaggttctcgaagccgagccgaagatgcgg 85
|||||
Db 182933 CAGAGGTTGTCGAGCCGAGCCGAGATGGCG 182799

RESULT 8
AC022708 68808 bp DNA HTG 24-MAR-2001
LOCUS Homo sapiens clone RP11-76H7, LOW-PASS SEQUENCE SAMPLING.
DEFINITION AC022708
ACCESSION AC022708.3 GI:13443223
VERSION HTG; HTGS_PHASE0.
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 68808)
AUTHORS Birren,B., Linton,L., Nusbaum,C. and Lander,E.
TITLE Homo sapiens chromosome, clone RP11-76H7
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 68808)
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
Anderson,S., Baldwin,J., Barna,N., Beckerly,R., Beda,F.,
Boguslavsky,L., Bouckgalter,B., Brown,A., Burkett,G., Castle,A.,
Choepey,Y., Colangelo,M., Collins,S., Collymore,A., Cooke,P.,
DeArrelano,K., Dewar,K., Domino,M., Doyle,M., Fenestor,J.,
Ferrelira,P., FitzHugh,W., Forrest,C., Gage,D., Galagan,J.,
Gardyna,S., Grant,G., Hagos,B., Heaford,A., Horton,L.,
Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J.,
Landers,T., Lehoczy,J., Levine,R., Lied,C., Liu,G., Locke,K.,
Macdonald,P., Marquis,N., McEwan,P., McGurk,A., McKernan,K.,
McPheeters,R., Meldrim,J., Meneus,L., Morrow,J., Naylor,J.,
Norman,C.H., O'Connor,T., O'Donnell,P., Oliver,T.M., Peterson,K.,
Pierre,N., Pisani,C., Pollara,V., Raymond,C., Riley,R., Rothman,D.,
Roy,A., Santos,R., Severy,P., Spencer,B., Stange-Thomann,N.,
Stojanovic,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
Tirrell,A., Vassiliev,H., Viel,R., Vo,A., Wu,X., Wyman,D., Ye,W.J.,
Zimmer,A. and Zody,M.
Direct Submission
Submitted (06-FEB-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Mar 24, 2001 this sequence version replaced gi:9146492.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L5354
Center clone name: 76_H7
-----
* NOTE: This record contains 84 individual
* sequencing reads that have not been assembled into
* contigs. Runs of N are used to separate the reads
* and the order in which they appear is completely

```

```

* arbitrary. Low-pass sequence sampling is useful for
* identifying clones that may be gene-rich and allows
* overlap relationships among clones to be deduced.
* However, it should not be assumed that this clone
* will be sequenced to completion. In the event that
* the record is updated, the accession number will
* be preserved.
* 1 717: contig of 717 bp in length
* 718 817: gap of 100 bp
* 818 1550: contig of 733 bp in length
* 1551 1650: gap of 100 bp
* 1651 2358: contig of 708 bp in length
* 2359 2458: gap of 100 bp
* 2459 3179: contig of 721 bp in length
* 3180 3279: gap of 100 bp
* 3280 4009: contig of 730 bp in length
* 4010 4109: gap of 100 bp
* 4110 4816: contig of 707 bp in length
* 4817 4916: gap of 100 bp
* 4917 5635: contig of 719 bp in length
* 5636 5735: gap of 100 bp
* 5736 6456: contig of 721 bp in length
* 6457 6556: gap of 100 bp
* 6557 7286: contig of 730 bp in length
* 7287 7386: gap of 100 bp
* 7387 8118: contig of 732 bp in length
* 8119 8218: gap of 100 bp
* 8219 8933: contig of 715 bp in length
* 8934 9033: gap of 100 bp
* 9034 9769: contig of 736 bp in length
* 9770 9869: gap of 100 bp
* 9870 10593: contig of 724 bp in length
* 10594 10693: gap of 100 bp
* 10694 11426: contig of 733 bp in length
* 11427 11526: gap of 100 bp
* 11527 12238: contig of 712 bp in length
* 12239 12338: gap of 100 bp
* 12339 13050: contig of 712 bp in length
* 13051 13150: gap of 100 bp
* 13151 13868: contig of 718 bp in length
* 13869 13968: gap of 100 bp
* 13969 14699: contig of 731 bp in length
* 14700 14799: gap of 100 bp
* 14800 15512: contig of 713 bp in length
* 15513 15612: gap of 100 bp
* 15613 16336: contig of 724 bp in length
* 16337 16436: gap of 100 bp
* 16437 17180: contig of 744 bp in length
* 17181 17280: gap of 100 bp
* 17281 17986: contig of 706 bp in length
* 17987 18086: gap of 100 bp
* 18087 18800: contig of 714 bp in length
* 18801 18900: gap of 100 bp
* 18901 19628: contig of 728 bp in length
* 19629 19728: gap of 100 bp
* 19729 20442: contig of 714 bp in length
* 20443 20542: gap of 100 bp
* 20543 21248: contig of 706 bp in length
* 21249 21348: gap of 100 bp
* 21349 22075: contig of 727 bp in length
* 22076 22175: gap of 100 bp
* 22176 22898: contig of 723 bp in length
* 22899 22998: gap of 100 bp
* 22999 23723: contig of 725 bp in length
* 23724 23823: gap of 100 bp
* 23824 24543: contig of 720 bp in length
* 24544 24643: gap of 100 bp
* 24644 25430: contig of 787 bp in length
* 25431 25530: gap of 100 bp
* 25531 26251: contig of 721 bp in length
* 26252 26351: gap of 100 bp
* 26352 27071: contig of 720 bp in length
* 27072 27171: gap of 100 bp

```

```

* 27172 27914: contig of 743 bp in length
* 27915 28014: gap of 100 bp
* 28015 28752: contig of 738 bp in length
* 28753 28852: gap of 100 bp
* 28853 28953: contig of 743 bp in length
* 29596 29695: gap of 100 bp
* 29696 30411: contig of 716 bp in length
* 30412 30511: gap of 100 bp
* 30512 31220: contig of 709 bp in length
* 31221 31320: gap of 100 bp
* 31321 32045: contig of 725 bp in length
* 32046 32145: gap of 100 bp
* 32146 32844: contig of 699 bp in length
* 32845 32944: gap of 100 bp
* 32945 33666: contig of 722 bp in length
* 33667 33766: gap of 100 bp
* 33767 34482: contig of 716 bp in length
* 34483 34582: gap of 100 bp
* 34583 35296: contig of 714 bp in length
* 35297 35396: gap of 100 bp
* 35397 36123: contig of 727 bp in length
* 36124 36223: gap of 100 bp
* 36224 36961: contig of 738 bp in length
* 36962 37061: gap of 100 bp
* 37062 37774: contig of 713 bp in length
* 37775 37874: gap of 100 bp
* 37875 38613: contig of 739 bp in length
* 38614 38713: gap of 100 bp
* 38714 39444: contig of 731 bp in length
* 39445 39544: gap of 100 bp
* 39545 40261: contig of 717 bp in length
* 40262 40361: gap of 100 bp
* 40362 40956: contig of 595 bp in length
* 40957 41056: gap of 100 bp
* 41057 41781: contig of 725 bp in length
* 41782 41881: gap of 100 bp
* 41882 42618: contig of 737 bp in length
* 42619 42718: gap of 100 bp
* 42719 43428: contig of 711 bp in length
* 43430 43529: gap of 100 bp
* 43530 44251: contig of 722 bp in length
* 44252 44351: gap of 100 bp
* 44352 45088: contig of 737 bp in length
* 45089 45188: gap of 100 bp
* 45189 45877: contig of 889 bp in length
* 45878 45977: gap of 100 bp
* 45978 46700: contig of 723 bp in length
* 46701 46800: gap of 100 bp
* 46801 47516: contig of 716 bp in length
* 47517 47616: gap of 100 bp
* 47617 48342: contig of 726 bp in length
* 48343 48442: gap of 100 bp
* 48443 49069: contig of 827 bp in length
* 49070 49169: gap of 100 bp
* 49170 49930: contig of 761 bp in length
* 49931 50030: gap of 100 bp
* 50031 50743: contig of 713 bp in length
* 50744 50843: gap of 100 bp
* 50844 51562: contig of 719 bp in length
* 51563 51662: gap of 100 bp
* 51663 52387: contig of 725 bp in length
* 52388 52487: gap of 100 bp
* 52488 53211: contig of 724 bp in length
* 53212 53311: gap of 100 bp
* 53312 54029: contig of 718 bp in length
* 54030 54129: gap of 100 bp
* 54130 54868: contig of 739 bp in length
* 54869 54968: gap of 100 bp
* 54969 55685: contig of 717 bp in length
* 55686 55785: gap of 100 bp
* 55786 56508: contig of 723 bp in length
* 56509 56608: gap of 100 bp
* 56609 57328: contig of 720 bp in length

```

```

* 57329 57428: gap of 100 bp
* 57429 58166: contig of 738 bp in length
* 58167 58266: gap of 100 bp

Query Match      3.4%  Score 28;  DB 2;  Length 68808;
Best Local Similarity 100.0%;  Pred. No. 0.0028;
Matches 28;  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;

QY 797 cctgccctcggggcaaaaaa 824
DB 33922 COTGCCCTCGGGCAAAAAA 33949

RESULT 9
ACOL12566/6
LOCUS ACOL12566.3 bp DNA HTG 03-MAR-2000
DEFINITION Homo sapiens chromosome 4 clone RP11-38D12 map 4, WORKING DRAFT
SEQUENCE 2 ordered pieces.
ACOL12566
ACOL12566.3 GI:7143435
HTG: HTGS_PHASE2; HTGS_DRAFT.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 172631)
AUTHORS Birren, B., Linton, L., Nusbaum, C. and Lander, E.
TITLE Homo sapiens chromosome 4, clone RP11-38D12
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 172631)
AUTHORS Birren, B., Linton, L., Nusbaum, C., Lander, E., Allen, N., Anderson, M.,
Baldwin, J., Barna, N., Beckler, R., Boguslavsky, L., Boukhgalter, B.,
Brown, A., Castle, A., Colangelo, M., Collins, S., Collymore, A.,
Cooke, P., DeArelano, K., Dewar, K., Domino, M., Donelan, L., Doyle, M.,
Ferreira, P., FitzHugh, W., Forrest, C., Funke, R., Gage, D.,
Galagan, J., Gardyna, S., Grant, G., Hagos, B., Heaford, A., Horton, L.,
Howland, J.C., Johnson, R., Jones, C., Kann, L., Karatas, A., Klein, J.,
Lehoczky, J., Lieu, C., Locke, K., Macdonald, P., Marquis, N.,
McEwan, P., McGurk, A., McKernan, K., McLaughlin, J., Meldrum, J.,
Morrow, J., Naylor, J., Norman, C.H., O'Connor, T., O'Donnell, P.,
Peterson, K., Pollara, V., Riley, R., Roy, A., Santos, R., Severy, P.,
Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J.,
Tesfaye, S., Tirrell, A., Vassiliev, H., Vo, A., Wheeler, J., Wu, X.,
Wyman, D., Ye, W.J., Zimmer, A. and Zody, M.
Direct Submission
Submitted (30-OCT-1999) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Mar 2, 2000 this sequence version replaced gi:5514022.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence.submissions@genome.wi.mit.edu
----- Project Information
Center project name: L2008
Center clone name: 38_D12
----- Summary Statistics
Sequencing vector: M13; M77815; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 162651 bases at least Q40
Consensus quality: 169868 bases at least Q30
Consensus quality: 171998 bases at least Q20
Insert size: 215000; agarose-fp
Insert size: 172531; sum-of-contigs
Quality coverage: 5.7 in Q20 bases; agarose-fp
Quality coverage: 7.1 in Q20 bases; sum-of-contigs
-----
* NOTE: this is a 'working draft' sequence. It currently
* consists of 2 contigs. Gaps between the contigs

```

\* are represented as runs of N. The order of the pieces  
 \* is believed to be correct as given, however the sizes  
 \* of the gaps between them are based on estimates that have  
 \* been provided by the submitter.  
 \* This sequence will be replaced  
 \* by the finished sequence as soon as it is available and  
 \* the accession number will be preserved.  
 \* 1 78380: contig of 78380 bp in length  
 \* 78381 78480: gap of 100 bp  
 \* 78481 172631: contig of 94151 bp in length.

# FEATURES

source  
 1. .172631  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /chromosome="4"  
 /map="4"  
 /clone\_lib="RP11-38D12"  
 /clone="RP11-38D12"  
 /note="assembly\_fragment  
 clone\_end:SP6  
 vector\_side:left"  
 78481..172631  
 /note="assembly\_fragment  
 clone\_end:17  
 vector\_side:right"  
 52774 a 30937 c 31576 g 57244 t 100 others

# BASE COUNT

# ORIGIN

Query Match 3.4%; Score 28; DB 2; Length 172631;  
 Best Local Similarity 100.0%; Fred. No. 0.0024;  
 Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 797 cctgcctcggcgcaaaaaa824

Db 91239 CCTGCCCTCGGCAAAAAA91212

# RESULT 10

# AC022249

# LOCUS

# DEFINITION

# AC022249

# VERSION

# KEYWORDS

# SOURCE

# ORGANISM

# REFERENCE

# AUTHORS

# TITLE

# JOURNAL

# REFERENCE

# AUTHORS

# 1

# (bases 1 to 198491)

# Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N.,

# Anderson,S., Baldwin,J., Barna,N., Beckerly,R., Bada,F.,

# Boguslavsky,L., Boukhgalter,B., Brown,A., Burkett,G., Castle,A.,

# Chapel,Y., Colangelo,M., Collins,S., Collymore,A., Cooke,P.,

# Deaerellano,K., Dewar,K., Domino,M., Doyle,M., Fenestor,J.,

# Ferreira,P., FitzHugh,W., Forrest,C., Gage,D., Galagan,J.,

# Gardyna,S., Grant,G., Hagos,B., Heaford,A., Horton,L.,

# Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J.,

# Landers,T., Lehoczy,J., Levine,R., Liu,C., Liu,G., Locke,K.,

# Macdonald,P., Marquis,N., McEwan,P., McGurk,A., McKernan,K.,

# McPheeters,R., Meldrim,J., Menes,L., Morrow,J., Naylor,J.,

# Norman,C.H., O'Connor,T., O'Donnell,P., Oliver,T.M., Peterson,K.,

# Pierre,N., Pisan,C., Pollara,V., Raymond,C., Riley,R., Rothman,D.,

# Roy,A., Santos,R., Severy,P., Spencer,B., Stange-Thomann,N.,

# Stojanovic,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,

# Tirrell,A., Vassiliev,H., Viel,R., Vo,A., Wu,X., Wyman,D., Ye.W.J.,

# Zimmer,A. and Zody,M.

# TITLE

# JOURNAL

# COMMENT

Direct Submission  
 Submitted (27-JAN-2000) Whitehead Institute/MIT Center for Genome  
 Research, 320 Charles Street, Cambridge, MA 02141, USA  
 On Sep 9, 2000 this sequence version replaced gi:5850453.  
 All repeats were identified using RepeatMasker:  
 Smit, A.F.A. & Green, P. (1996-1997)  
 http://ftp.genome.washington.edu/RM/RepeatMasker.html  
 ----- Genome Center  
 Center: Whitehead Institute/ MIT Center for Genome Research  
 Center code: WIBR  
 Web site: http://www-seq.wi.mit.edu  
 Contact: sequence\_submissions@genome.wi.mit.edu  
 ----- Project Information  
 Center project name: L4451  
 Center Clone name: 24\_G\_14  
 ----- Summary Statistics  
 Sequencing vector: M13; M7815; 100% of reads  
 Chemistry: Dye-terminator Big Dye; 100% of reads  
 Assembly program: Phrap; version 0.960731  
 Consensus quality: 187097 bases at least Q40  
 Consensus quality: 192525 bases at least Q30  
 Consensus quality: 194606 bases at least Q20  
 Insert size: 196000; agarose-fp  
 Insert size: 196691; sum-of-contigs  
 Quality coverage: 4.4 in Q20 bases; agarose-fp  
 Quality coverage: 4.4 in Q20 bases; sum-of-contigs

-----  
 \* NOTE: This is a 'working draft' sequence. It currently  
 \* consists of 19 contigs. The true order of the pieces  
 \* is not known and their order in this sequence record is  
 \* arbitrary. Gaps between the contigs are represented as  
 \* runs of N, but the exact sizes of the gaps are unknown.  
 \* This record will be updated with the finished sequence  
 \* as soon as it is available and the accession number will  
 \* be preserved.

1 23961: contig of 23961 bp in length  
 23962 24061: gap of 100 bp  
 24062 50896: contig of 26835 bp in length  
 50897 50996: gap of 100 bp  
 50997 52300: contig of 1304 bp in length  
 52301 52400: gap of 100 bp  
 52401 54995: contig of 2595 bp in length  
 54996 55095: gap of 100 bp  
 55096 57603: contig of 2508 bp in length  
 57604 57703: gap of 100 bp  
 57704 62215: contig of 4512 bp in length  
 62216 62315: gap of 100 bp  
 62316 66599: contig of 4284 bp in length  
 66600 66699: gap of 100 bp  
 66700 72051: contig of 5352 bp in length  
 72052 72151: gap of 100 bp  
 72152 79334: contig of 7183 bp in length  
 79335 79434: gap of 100 bp  
 79435 87876: contig of 8442 bp in length  
 87877 87976: gap of 100 bp  
 87977 95869: contig of 7893 bp in length  
 95870 95969: gap of 100 bp  
 95970 105886: contig of 9917 bp in length  
 105887 105986: gap of 100 bp  
 105987 116314: contig of 10328 bp in length  
 116315 116414: gap of 100 bp  
 116415 127405: contig of 10991 bp in length  
 127406 127505: gap of 100 bp  
 127506 140893: contig of 13388 bp in length  
 140894 140993: gap of 100 bp  
 140994 159375: contig of 18382 bp in length  
 159376 159475: gap of 100 bp  
 159476 178521: contig of 19046 bp in length  
 178522 178621: gap of 100 bp  
 178622 197646: contig of 19025 bp in length  
 197647 197746: gap of 100 bp  
 197747 198491: contig of 745 bp in length.

Location/Qualifiers

# FEATURES

[illegible]

Search completed: February 12, 2002, 19:09:51  
Job time: 5286 sec









[illegible]



```

Db 515 CCCACCTACGCCCGAGCGCTCTCTACGACATCCTGACATCGACAGGAGATCAGCG 574
QY 587 ggaagcgagggtgcagatctgttaccagcaggtagcgcgc 726
Db 575 GGAGCGGAGGTGTCAGATCTGTGTACCGCAGGTAGCGCGCG 614

RESULT 6
AW245467/c 586 bp mRNA EST 07-JAN-2000
LOCUS 2822901.3prime NIH_MGC_7 Homo sapiens cDNA clone IMAGE:2822901 3',
DEFINITION mRNA sequence.
ACCESSION AW245467
VERSION AW245467.1 GI:6588460
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE NIH-MGC http://mgi.nci.nih.gov/
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Other_ESTs: 2822901.5prime
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: DCTD/DTT cDNA Library Preparation: Ling
Hong/Rubin Laboratory cDNA Library Arrayed by: The I.M.A.G.E.
Consortium (LLNL) DNA Sequencing by: Berkeley MGC sequencing
project clone distribution: MGC clone distribution information can
be found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html Base Calling / Quality
Scores: PHRED from University of Washington Genome Center. Vector
Trimming: cross match from University of Washington Genome Center
PRAP suite. Poly-T Identification: patMatch.pl from Berkeley
Drosophila Genome Project. University of Washington Genome Center:
http://www.genome.washington.edu/polyadenylation: Based upon the
presence of a XhoI site followed by a run of 14 or more T residues
at the beginning of the sequence, this cDNA insert was
polyadenylated.
Plate: LICM10 row: J column: 22
High quality sequence stop: 474.
FEATURES
source
1..586
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2822901"
/clone_lib="NIH_MGC_7"
/tissue_type="small cell carcinoma"
/cell_line="MGC3"
/lab_host="DH10B (phage-resistant)"
/note="Organ: lung; Vector: pOTB7; Site_1: XhoI; Site_2:
EcoRI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGCACGAG(G). Size-selected >500bp for average
insert size 1.8kb. Library constructed by Ling Hong in
the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."
```

BASE COUNT  
ORIGIN

```

Query Match 69.9%; Score 576; DB 10; Length 586;
Best Local Similarity 100.0%; Pred. NO. 1.2e-152;
Matches 576; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 232 gtggctcccaaccagcagcgggcgagtgatgtgtggccaagctggagaccctctgc 291
Db 586 GTGGCTCCCAACACCAGCAGCGGGCGGAGTATGTGTGGCCAGCTGGATGACCTCGTC 527
QY 292 aactgggcccgcggaggtctctgtgcccattgaccttgcctgctgcgcgtg 351
```

```

Db 526 AACTGGGCCCCCGAGATTCTCTGTGCCCCATGACCTTGGCCCTGCTGCGCCGTCG 467
QY 352 gagatgatgacatgagcagcaccocccgtacgacatgagcagcgttttggtgtcttcgcgc 411
Db 466 GAGATGATGACATGGCAGCAGCACCCTGCTAGCAGATGACACCGCTTGGCGTGTCTCCGC 407
QY 412 gccagcccgccagtcgagcgtcatgctggtgcccgcacactcaccacaagatggcc 471
Db 406 GCCAGCCCGCCCGAGTCCGAGCTGATGATCGTGGCCGGACACTCACCACAAAGATGGCC 347
QY 472 ccagcgtttcgaaggtctcagcaccagatgccggagcgcgcgtactgtgtctccatgggg 531
Db 346 CCAGCGCTTGCAGAGGTCTACGACCCAGATCCGAGCGCGGTACGTGGTCTCCATGGGG 287
QY 532 agctgcgccaacggagaggtactaccactattctactcgtgtgtgagggcctgcac 591
Db 286 AGCTGGCGCAACGAGGAGGCTACTACCACTATTCTCTACTCGGTGTGAGGGCTCGCAC 227
QY 592 cgcctgtgcccgtggacatctacatcccaggtgcccacttcagccagggccctgctc 651
Db 226 CGCATGTGTCGGGTGGACATCTACATCCAGGCTGCCACCTACGCCCGAGGCCCTGCTC 167
QY 652 taaggcatctctgcagctgcagaggaagatcaagcggagcggagggctgcagatctggtac 711
Db 166 TACGCATCTCTGCGAGTCCAGAGGAAGATCAAGCGGAGCGAGGCTGCAGATCTGGTAC 107
QY 712 cgcaggtagcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 771
Db 106 CGCAGGTAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 47
QY 772 tgtgtcccgtaggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 807
Db 46 TGTGTCCCGTAGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 11

RESULT 7
BF342354 804 bp mRNA EST 22-NOV-2000
LOCUS 602013145F1 NCI_CGAP_Brn54 Homo sapiens cDNA clone IMAGE:4148842
DEFINITION 5', mRNA sequence.
ACCESSION BF342354
VERSION BF342354.1 GI:11289355
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE NIH-MGC http://mgi.nci.nih.gov/
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: David N. Louis, M.D.
cDNA Library Preparation: Life Technologies, Inc.
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM9409 row: j column: 11
High quality sequence start: 2
High quality sequence stop: 774.
FEATURES
source
1..804
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4148842"
/clone_lib="NCI_CGAP_Brn54"
/tissue_type="glioblastoma with EGFR amplification"
/lab_host="DH10B (RI phage-resistant)"
/note="Organ: brain; Vector: PCMV-SPORT6; Site_1: NotI;
```



DEFINITION	502269415F1 NIH_MGC_84 Homo sapiens cDNA clone IMAGE:4357803 '5', mRNA sequence.
ACCESSION	BF968344
VERSION	BF968344.1
KEYWORDS	GI:12335559
SOURCE	EST.
ORGANISM	human.
REFERENCE	Homo sapiens
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
TITLE	1 (bases 1 to 936)
JOURNAL	NIH-MGC <a href="http://mgc.nci.nih.gov/">http://mgc.nci.nih.gov/</a> .
COMMENT	National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished (1999) Contact: Robert Strausberg, Ph.D. Email: <a href="mailto:cgapbs-remail.nih.gov">cgapbs-remail.nih.gov</a> Tissue procurement: ATCC cDNA library Preparation: Life Technologies, Inc. cDNA Sequencing Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <a href="http://image.llnl.gov">http://image.llnl.gov</a> plate: LLAM9994 row: m column: 04 High quality sequence stop: 704. location/Qualifiers 1. 936 /organism="Homo sapiens" /db_xref="taxon:9606" /clone="IMAGE:4357803" /clone_lib="NIH_MGC_84" /tissue_type="adrenal cortex carcinoma, cell line" /lab_host="DH10B (phage-resistant)" /note="Organ: adrenal gland; Vector: pCMV-SPORT6; Site_1: Not; Site_2: SalI; Cloned unidirectionally; oligo-dT primed. Average insert size 1.229 Kb. Library enriched for full-length clones and constructed by Life Technologies. Note: this is a NIH MGC Library."
FEATURES	
source	

Qy	510	gocgtacgttggtctccatcaggagcgccgcgaacggaggagggattaccatacttacctta	569
Db	460	GCCTACGTGGTCTCCATCGGAGCTGGCCACGAGGAGGCTACTACCACTATTCTTA	519
Qy	570	ctcggtggtgagggactcgaccacatcgctgccgtgacatcatccaggctgc	628
Db	520	CTCGGTGTTGAGGGGCTCGCACCGCATCGTCGCCGTGSACATCTACATCCAGGCTGCC	578
RESULT	10		
LOCUS	BE897893		
DEFINITION	BE897893	883 bp mRNA EST 20-OCT-2000	
	601440565F1 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:3925168 5',		
	mRNA sequence.		
ACCESSION	BE897893		
VERSION	BE897893.1	GI:10363813	
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
REFERENCE	1	(bases 1 to 883)	
AUTHORS	NIH-MGC http://mgc.nci.nih.gov/.		
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)		
JOURNAL	Unpublished (1999)		
COMMENT	Contact: Robert Strausberg, Ph.D. Email: cgapbs-re@mail.nih.gov Tissue Procurement: ATCC/DCTD/DTF CDNA Library Preparation: Life Technologies, Inc. CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <a href="http://image.llnl.gov">http://image.llnl.gov</a> Plate: LMAW964 row: b column: 17 High quality sequence stop: 726.		
FEATURES	Location/Qualifiers		
source	1..883		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/clone="IMAGE:3925168"		
	/clone_lib="NIH-MGC-72"		
	/tissue_type="melanotic melanoma"		
	/lab_host="DH10B (phage-resistant)"		
	/note="Organ: skin; Vector: pCMV-SPORT6; Site:1: NotI; Site:2: SalI; Cloned unidirectionally. Primer: Oligo dt. Average insert size 2 kb. Library constructed by Life Technologies."		
BASE COUNT	185 a 276 c 282 g 140 t		
ORIGIN			
Query Match	65.0%; Score 536; DB 11; Length 883;		
Best Local Similarity	100.0%; Pred. No. 1.5e-141;		
Matches	536; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		
Qy	93	agtcttgcttcgctgcggttcctcgatccttggctgcctcacgctggccgctgt	152
Db	1	AGCTCTGGCTGCGGGCTTCGGATCCTTGTCGTGCTCCAGCGTGGGCCGGCTGT	60
Qy	153	gcaggcacagatgtccatcacagcgtggcccacgatggcccagcagcaccagcctgc	212
Db	61	GCAGGCACAGAGTGTCCATCAGAGCGTGGCCACCAGATGGCCCAAGCAGCACCAGCGCTGC	120
Qy	213	cctgccaaaggccagcgttgctccccaaacccagcagcggggcgagtatgtggtgc	272
Db	121	CCTGCCAAAGGCCAGAGCGGTGGCTCCCAAACCCAGCAGCGGGGGGAGTATGTGGTGGC	180
Qy	273	caagcttgtagacctgcactggcccccgcgagttctctgtggccataacctcgg	332
Db	181	CAAGCTGGATACCTCTGTCACTGGGCCCGCGGAGTCTCTGTGGCCCCATGACCTTCGG	240
Qy	333	cttgccctactcgcgcgtgagaataatcacatgtagcagccccgcgtcacacatgacgc	392





following 5' adaptor: GGCACGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies).

```

BASE COUNT      99 a      204 c      184 g      99 t
ORIGIN
Query Match      64.8%      Score 534;      DB 11;      Length 586;
Best Local Similarity 99.8%;      Pred. No. 7.3e-141;
Matches 584;      Conservative 0;      Mismatches 1;      Indels 0;      Gaps 0;

```

```

Tissue Procurement: ATCC
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
cDNA sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
Plate: LLC#73 row: c column: 06
High quality sequence stop: 744.
Location/Qualifiers
1. .773
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4101725"
/clone_lib="NIH_MGC_19"
/tissue_type="neuroblastoma"
/lab_host="DH10B (phage-resistant)"
/note="Organ: brain; Vector: pOTB7; Site:1: XhoI; Site:
EcoRI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5',
adaptor: GCACGAG(G). Library constructed by Ling Ho
in the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(stratagene) and Superscript II RT (Life Technologies
Note: this is a NIH_MGC Library."

```

```

BASE COUNT      140 a      247 c      249 g      137 t
ORIGIN
Query Match      63.8%; Score 526; DB 11; Length 773;
Best Local Similarity 100.0%; Pred. No. 1e-138;
Matches 526; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 103 ctgcggggttcgcgataccttggtctgcgtccacagctggcccgcgctctgcaggaacga 162
DB 26 CTCGCGGGTTCGGATCCTTGGTCTGCTGCCAGCTGGCGCCGCTGTGCAGGACGGA 85
QY 163 ggTgtccatcagagcgtggccacacgatggcccaagcagcaccacagcctgcctgcacaa 222
DB 86 GGTGTCCATCAGACGCTGGCCACCGATGGCCCRACGACGCCAGCTGCCCTGCCRAAG 145
QY 223 gccagagcgtggctcccaaacccacccagcagccggggagtagtatggtggccaaagctggat 282
DB 146 GCCAGAGCCGTGGCTCCCAAAACCCAGCAGCGGGCGGAGTATGTGTGGCCAAAGTGGAT 205
QY 283 gacctcgtcaactggccgcgcgagtgctctctgtggcccatgacctggcctggcctgc 342
DB 206 GACCTCGTCACTGGCGCGCGCGGAGTTCCTGTGGCCCATGACCTTCGGCCTGGCCTGC 265
QY 343 tgcgcgtggagatgatgcacatggcagcaccocgcgtacgacatgacgcgtttggcgtg 402
DB 266 TGGCGCGTGGAGATGCATGGCAGCACCCCGGTACGACATGGACGCGCTTTGGCGTG 325
QY 403 gtcttcgcgcagccgcgcagtcgcgacgtcatgatcgtggcggacacatccaccac 462
DB 326 GTCTTCGCGCCAGCCCGCGCCAGCTCCGACGCTCATGATCGTGGCGCGCACATCCACCAAC 385
QY 463 agatggccccagcgtctgcgaagttctacgacagatgccggagcgcgcgtacgtgctc 522
DB 386 AAGATGGCCCCAGCGCTTCGGAAGGTCTAGCACAGATGCGGAGCGCGGCTACGTGGTTC 445
QY 523 tccatgggagctgcgcaacggagaggtactaccattcttacttcgtgtgagg 582
DB 446 TCCATGGGAGCTGGCCACGGAGGAGGTACTACCACATTTCCTACTCGTGTGTGAGG 505
QY 583 ggcTgcgcgcgcacgtgctccgtggacatctacatccacggctgcc 628
DB 506 GGCTGGCAGCGCATCGTGCCTGGACATCTACATCCAGCGCTGCC 551

RESULT 14
BG419411
LOCUS
DEFINITION
        BG419411          910 bp      mRNA
        602446102F1 NIH_MGC 14 Homo sapiens cDNA clone IMAGE:4584528 5'
        EST
        14-MAR-2001

```

```

ACCESSION      BG419411
VERSION        BG419411.1  GI:13325917
KEYWORDS       mRNA sequence.
SOURCE         human.
ORGANISM       Homo sapiens
REFERENCE      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS       Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE         1. (bases 1 to 910)
JOURNAL       NIH-MGC http://mgi.nci.nih.gov/.
COMMENT       National Institutes of Health, Mammalian Gene Collection (MGC)
              Unpublished (1999)
              Contact: Robert Strausberg, Ph.D.
              Email: cgabbs@mail.nih.gov
              Tissue Procurement: DCTD/DP
              cDNA Library Preparation: Ling Hong/Rubin Laboratory
              CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
              DNA sequencing by: Incyte Genomics, Inc.
              Clone distribution: MGC clone distribution information can be
              found through the I.M.A.G.E. Consortium/LLNL at:
              http://image.llnl.gov
              Plate: LLC1310 row: d column: 01
              High quality sequence stop: 702.
              Location/Qualifiers
                1..910
                  /organism="Homo sapiens"
                  /db_xref="taxon:9606"
                  /clone="IMAGE:4584528"
                  /clone_lib="NIH_MGC_14"
                  /tissue_type="renal cell adenocarcinoma"
                  /lab_host="DH10B (phage-resistant)"
                  /note="Organ: kidney; Vector: pOTB7; Site_1: XhoI; Site_2:
                  ECORI; CDNA made by oligo-dt priming. Directionally
                  cloned into ECORI/XhoI sites using the following 5'
                  adaptor: GGCACGAG(G). Size-selected >500bp for average
                  insert size 1.8kb. Library constructed by Ling Hong in
                  the laboratory of Gerald M. Rubin (University of
                  California, Berkeley) using ZAP-cDNA synthesis kit
                  (Stratagene) and Superscript II RT (Life Technologies)."
```

BASE COUNT 195 a 279 c 278 g 156 t 2 others

ORIGIN

```

Query Match      61.4%; Score 506; DB 11; Length 910;
Best Local Similarity 99.8%; Pred. No. 3.8e-133;
Matches 556; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 72 gccaagatggcggtgtgtcagctcctgctgcgcgcgcgttcctcgatccttgggtctgcg 131
Db 1 GGCACAGATGGCGGTGTGTGTCAGCTCTGCTGCGCTGCGGCTCCGGATCCTTGGTCTGCG 60

Qy 132 ctccagcgtggccggcgtgtgcagcagaggtgtccatcagagcgtggccaccgatgg 191
Db 61 CTCAGCGTGGCGCTGCTGTGAGGCAGCAGGAGTGTCCATCAGAGCGTGGCCACCGATGG 120

Qy 192 cccaagcagcaccagcctgcctcccaagccagcagcgcgtggctcccaaccagcag 251
Db 121 CCAAGCAGCACCAGCGCTGCTGCCAAGGCCAGAGCGTGGCTCCCAACCCAGCAG 180

Qy 252 ccggggcgagtatgtgtggcgaagtggatgacctcgtcaactggccgcgcgcgagttc 311
Db 181 CCGGGCGAGTATGTGTGGCGCAAGCTGGATGACCTCGTCAACTGGCGCGCGCGAGTTC 240

Qy 312 tctgtgcccataccttgagcctgctgctgcgcgtggagatgacatgacagc 371
Db 241 TCTGTGCCCATGACCTTCGCGCTGGCTGCTGCGCGTGGAGATGATGACATGGCAGC 300

Qy 372 acccgcgtacgacatgacacgcttttggcgtgggtcttccgcgcgcagcccgccagtc 431
Db 301 ACCCGCTACGACATGACCGCTTTGGCGTGTGCTTCCGCGCGCAGCGCCGCGCATCGCA 360

Qy 432 cgtcatatcgtggcggcacactaccacaaagatggccccagcgttcgcaaggtcta 491
Db 1
```

BASE COUNT 125 a 252 c 235 g 126 t

ORIGIN

```

Query Match      61.3%; Score 505; DB 11; Length 738;
Best Local Similarity 99.8%; Pred. No. 8.4e-133;
Matches 555; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 73 gccaagatggcggtgtgtcagctcctgctgcgcgcgcgttcctcgatccttgggtctgcg 132
Db 1 GCCAAGATGGCGGTGTGTGTCAGCTCCTGCGCTGCGGCTCCGGATCCTTGGTCTGCGC 60

Qy 133 tcacagcgtggcccgactgtgagcagcagcagtggtccatcagcagcgtgcccaccatgac 192
Db 61 TCCAGCGTGGCGCTGCTGTGTCAGGCAGCAGGAGTGTCCATCAGAGCGTGGCCCGCATGCG 120
```

FEATURES

source

1..738

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/clone="IMAGE:4638441"

/clone\_lib="NIH\_MGC\_18"

/tissue\_type="large cell carcinoma"

/lab\_host="DH10B (phage-resistant)"

/note="Organ: lung; Vector: pOTB7; Site\_1: XhoI; Site\_2: ECORI; CDNA made by oligo-dt priming. Directionally cloned into ECORI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH-MGC Library."

BASE COUNT 125 a 252 c 235 g 126 t

ORIGIN

```

Query Match      61.3%; Score 505; DB 11; Length 738;
Best Local Similarity 99.8%; Pred. No. 8.4e-133;
Matches 555; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 73 gccaagatggcggtgtgtcagctcctgctgcgcgcgcgttcctcgatccttgggtctgcg 132
Db 1 GCCAAGATGGCGGTGTGTGTCAGCTCCTGCGCTGCGGCTCCGGATCCTTGGTCTGCGC 60

Qy 133 tcacagcgtggcccgactgtgagcagcagcagtggtccatcagcagcgtgcccaccatgac 192
Db 61 TCCAGCGTGGCGCTGCTGTGTCAGGCAGCAGGAGTGTCCATCAGAGCGTGGCCCGCATGCG 120
```

```
QY 193 ccaagcagacacccagcctgcccctgccaagccagagccgtggtctcccaaacccagcagc 252
Db 121 CCAAGCAGACACCCAGCCTGCCCTGCCAAAGGCCAGAGCCGTGGCTCCCAAACCCAGCAGC 180
QY 253 cggggcagtgatgtggtggccaaagtggatgacctgtaactggccgcccggaggttct 312
Db 181 CGGGCGAGTATGTGTGGCCAAAGCTGGATGACCTGCTCAACTGGGCCCCGCCGAGATTCT 240
QY 313 ctgtggcccatgaccttggccttgccctggtgcccgtggagatgatgcacatggcagca 372
Db 241 CTGTGGCCCATGACCTTCGGCCTGGCCTGGCCTGGCCTGGAGATGATGCACATGGCAGCA 300
QY 373 cccgctacgacatggaccactttggcgtggtcttccgcgcacccgcgcacgtccgac 432
Db 301 CCCGCTACGACATGGACCGCTTTGGCGTGGTCTTCGCCGCCAGCCGCCAGTCCGAC 360
QY 433 gtcattgatctggccggcacaactcaccacaagaatggccccagcgtttcgcaaggtctac 492
Db 361 GTCATGATCGTGGCCGGGCACACTCACCACAAGATGGCCGCCAGCGCTTCGCAAGGTCTAC 420
QY 493 gaccagatgccggagccgcgcgtacgtggtctccatggggagctgcgccaacggaggagc 552
Db 421 GACCAGATGCCGAGCCCGCGCTACGTGGTCTCCATGGGGAGCTGCCGCCAACCGGAGGAGC 480
QY 553 tactaccactattcctactcgtggtgaggggctggaccgcacatcgtgcccggtggacatc 612
Db 481 TACTACCACATATTCCTACTCGGTGCTGAGGGGCTGCCGCCGCATCGTGCCCGGTGGACATC 540
QY 613 tacatcccaggctgcc 628
Db 541 TACATCCGAGGCTGCC 556
```

Search completed: February 12, 2002, 18:08:53  
Job time: 2428 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 12, 2002, 20:29:55 ; Search time 33.73 Seconds  
(without alignments)  
481.031 Million cell updates/sec

Title: US-09-525-867-1  
Perfect score: 1118  
Sequence: 1 MAVLSAPGLRGFRILGRSS.....ILQLQKIKRRRIQIWR 213

Scoring table: BLOSOM62  
Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues  
Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR\_68:\*  
1: pir1.\*  
2: pir2.\*  
3: pir3.\*  
4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	965.5	86.4	216	2 S22371	NADH dehydrogenase
2	736	65.8	215	2 S48826	NADH dehydrogenase
3	732.5	65.5	213	2 T07603	NADH dehydrogenase
4	731	65.4	213	2 T44818	NADH dehydrogenase
5	726.5	63.0	199	2 T26329	hypothetical prote
6	725	64.8	218	2 S22286	NADH dehydrogenase
7	718	64.2	182	2 S78166	NADH dehydrogenase
8	654.5	58.5	174	2 F71692	NADH dehydrogenase
9	654	58.5	191	2 D82821	NADH-ubiquinone ox
10	626	56.0	173	2 E42573	psbG homolog prote
11	613	54.8	160	2 H81221	NADH dehydrogenase
12	605	54.1	160	2 F81922	NADH dehydrogenase
13	561	50.2	156	2 F3PPG	psbG protein - Par
14	521	46.6	179	2 C70413	NADH dehydrogenase
15	497.5	44.5	159	2 E64677	NADH dehydrogenase
16	496.5	44.4	159	2 C71838	NADH dehydrogenase
17	474	42.4	181	2 C75368	NADH dehydrogenase
18	461.5	41.3	167	2 H81252	NADH dehydrogenase
19	457	40.9	181	2 T11899	NADH dehydrogenase
20	444	39.7	220	1 E65000	NADH dehydrogenase
21	444	39.7	220	2 D85869	NADH dehydrogenase
22	443	39.6	184	2 T34623	NADH dehydrogenase
23	441.5	39.5	235	2 F84221	NADH dehydrogenase
24	438	39.2	248	1 S04437	NADH dehydrogenase
25	437.5	39.1	184	2 C70647	probable nuob prot
26	432	38.6	225	2 C83316	NADH dehydrogenase
27	431.5	38.6	247	1 F2NTG	psbG protein - com
28	428	38.3	224	2 A84948	NADH dehydrogenase
29	427.5	38.2	246	1 F2R2G	psbG protein - ric

30	426.5	38.1	243	1 F2LVG	psbG protein - liv
31	424.5	38.0	219	1 F2YB92	hypothetical ndhk
32	423.5	37.9	226	2 T09637	NADH dehydrogenase
33	423.5	37.9	245	2 S09666	psbG protein - whe
34	422.5	37.8	248	1 F2ZMG	NADH dehydrogenase
35	420.5	37.6	245	2 S74216	NADH dehydrogenase
36	408.5	36.5	224	2 S27978	NADH dehydrogenase
37	350.5	31.4	202	2 C72621	probable NADH dehy
38	344	30.8	96	2 FX0053	NADH dehydrogenase
39	319.5	28.6	353	2 D69478	NADH dehydrogenase
40	311	27.8	195	2 A75114	NADH dehydrogenase
41	301	26.9	195	2 D71019	probable NADH-ubiq
42	295	26.4	178	2 D72281	hypothetical prote
43	283	25.3	173	2 E71017	probable NADH-ubiq
44	281	25.1	170	2 D75115	co-induced hydroge
45	275	24.6	255	2 S08625	hydrogenase (EC 1.

ALIGNMENTS

RESULT 1

S22371  
NADH dehydrogenase (ubiquinone) (EC 1.6.5.3) chain PSST - bovine  
C:Species: Bos primigenius taurus (cattle)  
C:Date: 22-Nov-1993 #sequence\_revision 01-Dec-1995 #text\_change 26-Aug-1999  
C:Accession: S22371  
R:Arizmendi, J.M.; Runswick, M.J.; Skehel, J.M.; Walker, J.E.  
FEBS Lett. 301, 237-242, 1992  
A:Title: NADH: ubiquinone oxidoreductase from bovine heart mitochondria. A fourth ni  
A:Reference number: S22371; MUID:92249573  
A:Accession: S22371  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-216 <ARL>  
A:Cross-references: EMBL:X65020; NID:gl1256; PIDN:CAA46154.1; PID:9599691  
C:Superfamily: psbG protein  
C:Keywords: NAD; oxidoreductase

Query Match		86.4%	Score 965.5; DB 2; Length 216;
Best Local Similarity		85.7%	Pred. No. 2.6e-81;
Matches 186; Conservative 11; Mismatches 15; Indels 5; Gaps 2;			
QY	1	MAVLSAPGLRGFRILGRSSVGPVAVQARGVHQSVATDGPSTQPALPKARAVAPK----	P 56
Db	1	MAALAALRL-HPILAVRSGVGAALQVRGVHSSMAADSPSTQPAVQARAVVEKPAALP	59
QY	57	SRGEYVYVAKLDDLNVWARRSSLWPMTFGLACCAVEMHMAAPRYDMDRGVVFASPRQ	116
Db	60	SRGEYVYVAKLDDLNVWARRSSLWPMTFGLACCAVEMHMAAPRYDMDRGVVFASPRQ	119
QY	117	SDVMIVAGTLTKMAPALRKVYDQMPPEPVYVSMGSCANGGGYHYSVYVVGCDRIYVP	176
Db	120	SDVMIVAGTLTKMAPALRKVYDQMPPEPVYVSMGSCANGGGYHYSVYVVGCDRIYVP	179
QY	177	DIYFGCPPTAEALLYGILQKRIKRRRLQIWR 213	
Db	180	DIYFGCPPTAEALLYGILQKRIKRRRLQIWR 216	

RESULT 2

S48826  
NADH dehydrogenase (ubiquinone) (EC 1.6.5.3) chain PSST precursor - wild cabbage  
C:Species: Brassica oleracea (wild cabbage)  
C:Date: 13-Jan-1995 #sequence\_revision 13-Jan-1995 #text\_change 23-Mar-2001  
C:Accession: S48826  
R:Popson, B.J.; Downs, C.G.; Davies, K.M.; Morris, S.C.; Buchanan-Wollaston, V.  
submitted to the EMBL Data Library, September 1994  
A:Description: Nucleotide sequence of a cDNA clone from broccoli (Brassica oleracea  
A:Reference number: S48827  
A:Accession: S48826  
A:Molecule type: mRNA



Db 10 AVGTERRLASTQAIAGNSEA-PKGIATTGTPTLPNPSKAEYALARDLDDVLNLAQRGSWPL 68  
 QY 83 TFLGACCAVEMHMAAPRYDMDRFQWFRASPROSDVMIVAGTUTNKMAPALRKVDYDMP 142  
 Db 69 TFLGACCAVEMHMAAPRYDMDRFQWFRASPRODLIFVAGTUTNKMAPALRIIDMP 128  
 QY 143 EPRVYVSMGSCANGGYYHYHSYVVRGCDRIVPVDIIPGCPPTAEALLYGILOLQKIK 202  
 Db 129 EAKWISMGSCANGGYYHYHSYVLRGCDRIVPVDIIPGCPPTAEALLYGILOLQKIK 188  
 QY 203 RERRLQIYWR 213  
 Db 189 RXREAQLWRR 199  
 RESULT 6  
 S52286  
 NADH dehydrogenase (ubiquinone) (EC 1.6.5.3) chain PSST precursor - Arabidopsis thaliana  
 N:Alternate names: protein T22P22.160  
 C:Species: Arabidopsis thaliana (mouse-ear cress)  
 C:Date: 08-May-1995 #sequence\_revision 10-Nov-1995 #text\_change 19-May-2000  
 C:Accession: S52286; T48536  
 R:Heiser, V.; Grohmann, L.  
 submitted to the EMBL Data Library, January 1995  
 A:Reference number: S52286  
 A:Accession: S52286  
 A:Molecule type: mRNA  
 A:Residues: 1-218 <HE>  
 A:Cross-references: EMBL:X84078; NID:g643089; PIDN:CAA58887.1; PID:g643090  
 R:Bevan, M.; Hilbert, H.; Braun, M.; Holzer, E.; Brandt, A.; Duesterhoeft, A.; Bancroft,  
 submitted to the Protein Sequence Database, April 2000  
 A:Reference number: Z24490  
 A:Accession: T48536  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-218 <BE>  
 A:Cross-references: EMBL:AL163814  
 A:Experimental source: cultivar Columbia; BAC clone T22P22  
 C:Genetics:  
 A:Gene: PSST  
 A:Map position: 5  
 A:Genome: nuclear  
 A:Introns: 141/2  
 A:Note: T22P22.160  
 C:Superfamily: psbg protein  
 C:Keywords: mitochondrion; NAD; oxidoreductase

Query Match 64.8%; Score 725; DB 2; Length 218;  
 Best Local Similarity 68.2%; Pred. No. 3.1e-59;  
 Matches 135; Conservative 20; Mismatches 35; Indels 8; Gaps 2;  
 QY 24 AVOARGV---HQSATDCGSS-----TOPALPKARAVAPKPSRGEVYVAKLDDLNNWR 75  
 Db 21 AVAAASVSHLTSLPALSPSTSPSTSPSTSPSTSPSTSPSTSPSTSPSTSPSTSPSTSPST 80  
 QY 76 RSLPMTFGLACCAVEMHMAAPRYDMDRFQWFRASPROSDVMIVAGTUTNKMAPALR 135  
 Db 81 TGSIPMTFGLACCAVEMHMAAPRYDMDRFQWFRASPROSDVMIVAGTUTNKMAPALR 140  
 QY 136 KYVDQMPPEPRVYVSMGSCANGGYYHYHSYVVRGCDRIVPVDIIPGCPPTAEALLYGI 195  
 Db 141 KYVDQMPPEPRVYVSMGSCANGGYYHYHSYVVRGCDRIVPVDIIPGCPPTAEALLYGL 200  
 QY 196 QLQRIKRRRLQIYWR 213  
 Db 201 QLQKINRRKDFLHWNK 218  
 RESULT 7  
 S78166  
 NADH dehydrogenase (ubiquinone) (EC 1.6.5.3) chain 10 - Reclinomonas americana (ATCC 503  
 C:Species: mitochondrion Reclinomonas americana

A:Variety: ATCC 50394  
 C:Date: 29-Jan-1998 #sequence\_revision 06-Feb-1998 #text\_change 26-Aug-1999  
 C:Accession: S78166  
 R:Lang, B.F.; Burger, G.; O'Kelly, C.J.; Cedergren, R.; Golding, G.B.; Lemieux, C.;  
 Nature 387, 493-497, 1997  
 A:Title: An ancestral mitochondrial DNA resembling a eubacterial genome in miniatur.  
 A:Reference number: S78127; MUID:97311393  
 A:Accession: S78166  
 A:Status: nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-182 <LAN>  
 A:Cross-references: EMBL:AF007261; NID:g2258325; PIDN:AAD11899.1; PID:g2258365  
 A:Experimental source: ATCC 50394  
 A:Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1997  
 C:Genetics:  
 A:Gene: nad10  
 A:Genome: mitochondrion  
 C:Superfamily: psbg protein  
 C:Keywords: mitochondrion; NAD; oxidoreductase

Query Match 64.2%; Score 718; DB 2; Length 182;  
 Best Local Similarity 81.4%; Pred. No. 1.1e-58;  
 Matches 127; Conservative 18; Mismatches 11; Indels 0; Gaps 0;  
 QY 58 SRGEYVVAKLDDLNNWARRSSLNPMFTFGLACCAVEMHMAAPRYDMDRFQWFRASPROS 117  
 Db 27 NKTETIVSKMDLNNWARRKSLNPMFTFGLACCAVEMHMAAPRYDMDRFQWFRASPROS 86  
 QY 118 DVMIVAGTUTNKMAPALRKVDQMPPEPRVYVSMGSCANGGYYHYHSYVVRGCDRIVPVD 177  
 Db 87 DVMIVAGTUTNKMAPALRKVDQMPPEPRVYVSMGSCANGGYYHYHSYVVRGCDRIVPVD 146  
 QY 178 IYVPCPTAEALLYGILOLQRIKRRRLQIYWR 213  
 Db 147 IYVPCPTAEALLYGILOLQRIKRRRLQIYWR 182  
 RESULT 8  
 F71692  
 NADH dehydrogenase (ubiquinone) (EC 1.6.5.3) I chain B RP356 - Rickettsia prowazeki  
 C:Species: Rickettsia prowazekii  
 C:Date: 21-Nov-1998 #sequence\_revision 21-Nov-1998 #text\_change 03-Nov-2000  
 C:Accession: F71692  
 R:Andersson, S.G.E.; Zomorodipour, A.; Andersson, J.O.; Sicheritz-Ponten, T.; Alsm  
 Nature 396, 133-140, 1998  
 A:Title: The genome sequence of Rickettsia prowazekii and the origin of mitochondri.  
 A:Reference number: A71630; MUID:99039499  
 A:Accession: F71692  
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-174 <AND>  
 A:Cross-references: GB:A7235271; GB:A7235269; NID:g3868717; PIDN:CAA14816.1; PID:g3  
 A:Experimental source: strain Madrid E  
 C:Genetics:  
 A:Gene: nuob; RP356  
 C:Superfamily: psbg protein  
 C:Keywords: NAD; oxidoreductase

Query Match 58.5%; Score 654.5; DB 2; Length 174;  
 Best Local Similarity 78.2%; Pred. No. 7.1e-53;  
 Matches 115; Conservative 25; Mismatches 6; Indels 1; Gaps 1;  
 QY 57 SRGEYVVAKLDDLNNWARRSSLNPMFTFGLACCAVEMHMAAPRYDMDRFQWFRASPRO 116  
 Db 17 TNRG-FLLSKYDEVISWARRSLNPMFTFGLACCAVEMHMAAPRYDMDRFQWFRASPRO 75  
 QY 117 SDVMIVAGTUTNKMAPALRKVDQMPPEPRVYVSMGSCANGGYYHYHSYVVRGCDRIVP 176  
 Db 76 SDLMIVAGTUTNKMAPALRKVDQMPPEPRVYVSMGSCANGGYYHYHSYVVRGCDRIVP 135  
 QY 177 DIYIPGCPPTAEALLYGILOLQRIK 203





A:Reference number: A81775; MUID:20222556  
A:Accession: F81992  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-160 <PAR>  
A:Cross-references: GB:AL162752; GB:AL157959; NID:g7378778; PIDN:CAB83338.1; PID:g737879  
A:Experimental source: serogroup A, strain Z2491  
C:Genetics:  
A:Gene: nucB; NMA0018  
C:Superfamily: psbG protein  
C:Keywords: NAD; oxidoreductase

Query Match 54.1%; Score 605; DB 2; Length 160;  
Best Local Similarity 73.2%; Pred. No. 2.3e-48;  
Matches 104; Conservative 22; Mismatches 16; Indels 0; Gaps 0;

Qy 62 YVAKLDDLNNARRSLWPTFGLACCAVEMHMAAPRYDMDFGVFRASPROSDVMI 121  
Db 11 FITTSADTVLNYMTGSLWPTFGLACCAVEMHMAAPRYDMDFGVFRASPROADLMI 70  
Qy 122 VAGTLTNKMAPALRKVYDQMPPEPRVYVSMGSCANGGGYHYHSYVVRGCDRIVPVDIKIP 181  
Db 71 VAGTLTNKMAPALRKVYDQMPPEPRVYVSMGSCANGGGYHYHSYVVRGCDRIVPVDIKIP 130  
Qy 182 GCPPTAEALYGLIQLOKIKR 203  
Db 131 GCPPTAEALYGLIQLOKIKR 152

RESULT 13  
F2PBG  
psbG protein - Paramesium tetraurelia mitochondrion  
C:Species: mitochondrion Paramesium tetraurelia  
C:Date: 30-Sep-1991 #sequence\_revision 30-Sep-1991 #text\_change 19-Jan-2001  
C:Accession: S07736; JS0235  
R:Pritchard, A.E.; Seilhamer, J.J.; Mahalingam, R.; Sable, C.L.; Venuti, S.E.; Cummings, Nucleic Acids Res. 18, 173-180, 1990  
A:Title: Nucleotide sequence of the mitochondrial genome of Paramesium.  
A:Reference number: S07725; MUID:90174913  
A:Accession: S07736  
A:Status: translation not shown  
A:Molecule type: DNA  
A:Residues: 1-156 <PRI>  
A:Cross-references: EMBL:X15917; NID:g13256; PIDN:CAA34045.1; PID:g578753  
R:Pritchard, A.E.; Venuti, S.E.; Ghalambor, M.A.; Sable, C.L.; Cummings, D.J. Gene 78, 121-134, 1989  
A:Title: An unusual region of Paramesium mitochondrial DNA containing chloroplast-like  
A:Reference number: JS0231; MUID:89357489  
A:Accession: JS0235  
A:Molecule type: DNA  
A:Residues: 1-122, '126', '128-156 <PRI2>  
A:Cross-references: GB:M26930; NID:g341550; PIDN:AAA79257.1; PID:g1019632  
A:Experimental source: strain sp. 4.51  
A:Note: The source is designated as Paramesium aurelia species 4 stock 51, now designated  
C:Genetics:  
A:Gene: psbG  
A:Genome: mitochondrion  
A:Genetic code: Sec6  
A:Start codon: ATA  
C:Superfamily: psbG protein  
C:Keywords: mitochondrion

Query Match 50.2%; Score 561; DB 1; Length 156;  
Best Local Similarity 63.4%; Pred. No. 2.4e-44;  
Matches 92; Conservative 33; Mismatches 20; Indels 0; Gaps 0;

Qy 59 RGEYVAKDDLNNARRSLWPTFGLACCAVEMHMAAPRYDMDFGVFRASPROSD 118  
Db 4 KADFLKLSANNLSHWKQSGFWLTLGLACCALEMMHAIYSRYDFRFGVIFRATPRQAD 63  
Qy 119 VMVAGTLTNKMAPALRKVYDQMPPEPRVYVSMGSCANGGGYHYHSYVVRGCDRIVPVDI 178

Db 64 LIIVAGTVTNKMAPALRRLYDQTADPKWLWSMGSCANGGGYHYSAVYKGCCKIIPVDM 123  
Qy 179 YIPGCPPTAEALYGLIQLOKIKR 203  
Db 124 LCPRCPTAEALFFGVQLQLOKTLMK 148  
RESULT 14  
C70413  
NADH dehydrogenase (ubiquinone) (EC 1.6.5.3) I chain nuoB - Aquifex aeolicus  
C:Species: Aquifex aeolicus  
C:Date: 08-May-1998 #sequence\_revision 08-May-1998 #text\_change 05-Nov-1999  
C:Accession: C70413  
R:Deckert, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.  
V: Nature 392, 353-358, 1998  
A:Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.  
A:Reference number: A70300; MUID:98196666  
A:Accession: C70413  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-179 <AQF>  
A:Cross-references: GB:AE000734; NID:g2983733; PIDN:AAC07297.1; PID:g2983736; GB:AE000734  
A:Experimental source: strain VF5  
C:Genetics:  
A:Gene: nuoB  
C:Superfamily: psbG protein  
C:Keywords: NAD; oxidoreductase

Query Match 46.6%; Score 521; DB 2; Length 179;  
Best Local Similarity 60.8%; Pred. No. 1.3e-40;  
Matches 90; Conservative 33; Mismatches 23; Indels 2; Gaps 2;

Qy 57 SSRGEYVAKDDLNNARRSLWPTFGLACCAVEMHMAAPRYDMDFGVFRASPRO 116  
Db 5 NSNG-FVTITVELLRNGRNSLWPTFGLACCAIEMHTAASRFDLRLGVIFRASPRO 63  
Qy 117 SDVMIVAGTLTNKMAPALRKVYDQMPPEPRVYVSMGSCANGGGYHYHSYVVRGCDRIVPV 176  
Db 64 ADVLIVAGTVTNKMAPALRKVYDQMPPEPRVYVSMGSCANGGGYHYHSYVVRGCDRIVPV 122  
Qy 177 DIYIPGCPPTAEALYGLIQLOKIKR 204  
Db 123 DVIYIPGCPPTFQGLIYGLIQLOKIKR 150  
RESULT 15  
E64677  
NADH dehydrogenase (ubiquinone) (EC 1.6.5.3) chain NQ06 - Helicobacter pylori (strai  
C:Species: Helicobacter pylori  
C:Date: 09-Aug-1997 #sequence\_revision 09-Aug-1997 #text\_change 26-Aug-1999  
C:Accession: E64677  
R:Tomb, J.F.; White, O.; Kerlavage, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann, Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.; Khalak, H.G.; Glodek, A.; McKenney, J.D.; Kelley, J.M.; Cotton, M.D.; Weidman, J.M.; Fujii, C.; Bowman, C.; Watthey Nature 388, 539-547, 1997  
A:Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karpk, P.D.; Smith, H.O.; Fraser A:Title: The complete genome sequence of the gastric pathogen Helicobacter pylori.  
A:Reference number: A64520; MUID:97394467  
A:Accession: E64677  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-159 <TOM>  
A:Cross-references: GB:AE000631; GB:AE000511; NID:g2314421; PIDN:AAD08307.1; PID:g2314421  
C:Superfamily: psbG protein  
C:Keywords: electron transfer; membrane-associated complex; NAD; oxidoreductase

Query Match 44.5%; Score 497.5; DB 2; Length 159;  
Best Local Similarity 59.6%; Pred. No. 1.7e-38;  
Matches 87; Conservative 29; Mismatches 29; Indels 1; Gaps 1;

QY 63 VVAKLDDLVNWARSSSLWPTMTGLACCAVEMHMAAPRYDMDRFGVVERASPROSDVMIV 122  
Db 7 VLSLTKLLNWGSNSLWPTTYGLACCAIEMMATGGSRFDFDRFGTIFRASPROSDVMII 66  
QY 123 AGTLTKKMAPALKYVDQMPPEPRYVVMGSCANGGYYHYYSVVRGCDRIVPVDIYIPG 182  
Db 67 AGTLTKKHAEFMRRLYDQMPPEPKWVISMGCANTGGMEN-TYATVOGADRVVPVDIVLPG 125  
QY 183 CPPTAEALLYGILQORKIKRERRLQ 208  
Db 126 CAPRPETLQYALMWLODKIRSKAIK 151

Search completed: february 12, 2002, 20:45:24  
Job time: 929 sec

Result No.	Score	Query Match	Length	DB	ID	Description
1	750.8	91.1	772	22	AAI58780	Human polynucleoti
2	738.8	89.7	752	22	AAI60566	Human polynucleoti
3	715.4	86.8	785	21	AAC59553	Human secreted pro
C 4	384.6	46.7	789	19	AAV61275	3' cDNA sequence o
C 5	384.6	46.7	789	19	AAV58509	3' fragment of pro
C 6	384.6	46.7	789	21	AAA06272	Human immunogenic
C 7	384.6	46.7	789	22	AAI10031	Human prostate tum
C 8	384.6	46.7	789	22	AAH93388	Human prostate-spe
C 9	384.6	46.7	789	22	AAH84702	Human prostate-spe
C 10	384.6	46.7	789	22	AAH02453	Prostate tumour an
C 11	279.2	33.9	911	21	AAC43639	zebra mays DNA fragm

(HYSE-) HYSEQ INC.

Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;  
Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;  
Zhao QA, Zhou P, Goodrich R, Drmanac RT;

WPI; 2001-442253/47.

P-PSDB; AAM39624.

Novel nucleic acids and polypeptides, useful for treating disorders

such as central nervous system injuries -

Claim 1; SEQ ID NO 983; 10078pp: English.

The invention relates to human nucleic acids (AAI57798-AAI61369) and the encoded polypeptides (AAM39642-AAAM42213) with neurotrophic, immunosuppressant and cytostatic activity. The polynucleotides are useful in gene therapy. A composition containing a polypeptide or polynucleotide of the invention may be used to treat diseases of the peripheral nervous system, such as peripheral nervous injuries, peripheral neuropathy and localised neuropathies and central nervous system diseases, such as Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic lateral sclerosis, and Shy-Drager Syndrome. Other uses include the utilisation of the activities such as: immune system suppression, Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic and thrombolytic activity, cancer diagnosis and therapy, drug screening, assays for receptor activity, arthritis and inflammation, leukaemias and C.N.S disorders.

Note: The sequence data for this patent did not form part of the printed specification.

Sequence 772 BP; 136 A; 265 C; 245 G; 126 T; 0 other:

Query Match	91.1%;	Score 750.8;	DB 22;	Length 772;
Best Local Similarity	99.1%;	Pred. No. 1.8e-155;		
Matches 755; Conservative	0;	Mismatches 7;	Indels 0;	Gaps 0;

Qy 63 gaaggccgagccaaagatggcggtgctgtcagctctctgacctgacgagcttccgqatcct 122

122

Qy	123	tggctgcgctccagcgtggggcccgctgtgcagqccacgagqatgtccatcagagcgtgc	182

[illegible]

Qy 183 caccgatagcccaagcagcacccagccctgccctgcccaagcccaagccctgccctcccaa 242

242

Ov 243 acccagccagccccgggcccgaattatcttgccttcgcccaacctaaccctccctcacctccgcccga

QY 243 acccagcagccggggcgagatagtgtggccaaagtggatgacctcgtcaactgggcccg 302

200 100 50 0

Qy 303 ccggaggtctctctgtggcccatgaccttcggcctggcctgctggccgctggagatgatgca 362

[illegible]

363 catggcagcaccgcgctacgacatggaccgcttggcgtggtcttcgcgcagccgcg 422  
QY

[illegible]

Qy 423 ccagtccgacgtcatgatcgtggccggcacactcaccaagaatggccccagcgcttcg 482

12345678910111213141516171819202122232425262728293031323334353637383940414243444546474849505152535455565758596061626364656667686970717273747576777879808182838485868788899091929394959697989910010110210310410510610710810911011111211311411511611711811912012112212312412512612712812913013113213313413513613713813914014114214314414514614714814915015115215315415515615715815916016116216316416516616716816917017117217317417517617717817918018118218318418518618718818919019119219319419519619719819920020120220320420520620720820921021121221321421521621721821922022122222322422522622722822923023123223323423523623723823924024124224324424524624724824925025125225325425525625725825926026126226326426526626726826927027127227327427527627727827928028128228328428528628728828929029129229329429529629729829930030130230330430530630730830931031131231331431531631731831932032132232332432532632732832933033133233333433533633733833934034134234334434534634734834935035135235335435535635735835936036136236336436536636736836937037137237337437537637737837938038138238338438538638738838939039139239339439539639739839940040140240340440540640740840941041141241341441541641741841942042142242342442542642742842943043143243343443543643743843944044144244344444544644744844945045145245345445545645745845946046146246346446546646746846947047147247347447547647747847948048148248348448548648748848949049149249349449549649749849950050150250350450550650750850951051151251351451551651751851952052152252352452552652752852953053153253353453553653753853954054154254354454554654754854955055155255355455555655755855956056156256356456556656756856957057157257357457557657757857958058158258358458558658758858959059159259359459559659759859960060160260360460560660760860961061161261361461561661761861962062162262362462562662762862963063163263363463563663763863964064164264364464564664764864965065165265365465565665765865966066166266366466566666766866967067167267367467567667767867968068168268368468568668768868969069169269369469569669769869970070170270370470570670770870971071171271371471571671771871972072172272372472572672772872973073173273373473573673773873974074174274374474574674774874975075175275375475575675775875976076176276376476576676776876977077177277377477577677777877978078178278378478578678778878979079179279379479579679779879980080180280380480580680780880981081181281381481581681781881982082182282382482582682782882983083183283383483583683783883984084184284384484584684784884985085185285385485585685785885986086186286386486586686786886987087187287387487587687787887988088188288388488588688788888989089189289389489589689789889990090190290390490590690790890991091191291391491591691791891992092192292392492592692792892993093193293393493593693793893994094194294394494594694794894995095195295395495595695795895996096196296396496596696796896997097197297397497597697797897998098198298398498598698798898999099199299399499599699799899910001001100210031004100510061007100810091010101110121013101410151016101710181019102010211022102310241025102610271028102910301031103210331034103510361037103810391040104110421043104410451046104710481049105010511052105310541055105610571058105910601061106210631064106510661067106810691070107110721073107410751076107710781079108010811082108310841085108610871088108910901091109210931094109510961097109810991100110111021103110411051106110711081109111011111112111311141115111611171118111911201121112211231124112511261127112811291130113111321133113411351136113711381139114011411142114311441145114611471148114911501151115211531154115511561157115811591160116111621163116411651166116711681169117011711172117311741175117611771178117911801181118211831184118511861187118811891190119111921193119411951196119711981199120012011202120312041205120612071208120912101211121212131214121512161217121812191220122112221223122412251226122712281229123012311232123312341235123612371238123912401241124212431244124512461247124812491250125112521253125412551256125712581259126012611262126312641265126612671268126912701271127212731274127512761277127812791280128112821283128412851286128712881289129012911292129312941295129612971298129913001

Qy 483 caaggctacgaccagatgccggagccgcgtacgtggtctccatggggagctgcgcaa 542

[illegible]

Qy 543 cggagggtactaccactattcctactcgggtggtgaggggctgcgaccgcgtgcc 602

CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic  
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the  
CC utilisation of the activities such as: Immune system suppression,  
CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic  
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,  
CC assays for receptor activity, arthritis and inflammation, leukaemias and  
CC C.N.S disorders  
CC Note: The sequence data for this patent did not form part of the printed  
CC specification.  
XX  
SQ Sequence 752 BP; 130 A; 260 C; 236 G; 126 T; 0 other;

Query Match 89.7%; Score 738.8; DB 22; Length 752;  
Best Local Similarity 99.1%; Pred. No. 7.4e-153;  
Matches 743; Conservative 0; Mismatches 7; Indels 0; Gaps 0;  
QY 75 caagatggcgggtgctgcagctcctggcctgcggcgttcggatccttggtgtgcgtc 134  
Db 1 caagatggcgggtgctgcagctcctggcctgcggcgttcggatccttggtgtgcgtc 60  
QY 135 cagcgtggcgcggcgtgctgcaggcaagaggtgtccatcagagcgtggccaccgatggccc 194  
Db 61 cagcgtggcgcggcgtgctgcaggcaagaggtgtccatcagagcgtggccaccgatggccc 120  
QY 195 aagcagcaccagcctgcctggccaaagccagagcgtgctcccaaaaccagcagccg 254  
Db 121 aagcagcaccagcctgcctggccaaagccagagcgtgctcccaaaaccagcagccg 180  
QY 255 gggcagatgattggtggccaaagctgagtagcctgcaactgggcccgcggaggtctct 314  
Db 181 gggcagatgattggtggccaaagctgagtagcctgcaactgggcccgcggaggtctct 240  
QY 315 qfsgcccatgacttcgcctgcctgcctgcctgcctgcctgcctgcctgcctgcctgcct 374  
Db 241 gfgcccatgacttcgcctgcctgcctgcctgcctgcctgcctgcctgcctgcctgcct 300  
QY 375 cgcgtacagatggaccgcttggcgtggttcttcgcgcagcgcgcgcgcgcgcgcgcgcgc 434  
Db 301 cgcgtacagatggaccgcttggcgtggttcttcgcgcagcgcgcgcgcgcgcgcgcgcgc 360  
QY 435 catgctgctggccgc 494  
Db 361 catgctgctggccgc 420  
QY 495 ccagatggcggagc 554  
Db 421 ccagatggcggagc 480  
QY 555 ctaccactattcctactcgtggtggtggtggtggtggtggtggtggtggtggtggtggtggt 614  
Db 481 ctaccactattcctactcgtggtggtggtggtggtggtggtggtggtggtggtggtggtggt 540  
QY 615 catccagcgtcccaactacgc 674  
Db 541 catccagcgtcccaactacgc 600  
QY 675 gaagatcaagcggagcggagcgtcagatctggtaccgcaggttagcgcgcgcgcgcgcgcgcgc 734  
Db 601 gaagatcaagcggagcggagcgtcagatctggtaccgcaggttagcgcgcgcgcgcgcgcgcgc 660  
QY 735 cgc 794  
Db 661 cgc 720  
QY 795 aacctgcccctcgggcaaaaaa 824  
Db 721 aacctgcccctcgggcrccgcaaaaaa 750

RESULT 3  
AAC59553  
ID AAC59553 standard; cDNA; 785 BP.

XX AAC59553;  
AC 15-FEB-2001 (first entry)  
DT Human secreted protein cDNA sequence #47.  
DE  
KW Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;  
KW anti-allergic; hepatotropic; antidiabetic; antinflammatory; antitumor;  
KW vulnary; anticonvulsant; antibacterial; antifungal; antiparasitic;  
KW cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;  
KW neurological disease; infection; human; secreted protein; ss.  
XX  
OS Homo sapiens.  
PN WO200055352-A2.  
XX  
XX 21-SEP-2000.  
XX 09-MAR-2000; 2000WO-US06044.  
XX 12-MAR-1999; 99US-0124099.  
XX 03-DEC-1999; 99US-0188661.  
XX (HUMA-) HUMAN GENOME SCI INC.  
XX PA  
XX Rosen CA, Ruben SM, Komatsoulis G;  
XX  
XX WPI; 2000-602124/57.  
XX P-PSDB; AAB34261.  
XX  
XX Novel human secreted proteins useful for diagnosis, prevention and  
XX treatment of disorders including neurological, cell proliferative,  
XX cardiovascular, autoimmune and inflammatory disorders and microbial  
XX infections  
XX  
XX Claim 1; Page 345; 383pp; English.  
XX The invention relates to the isolation of genes AAC59507-C59556 encoding  
XX 50 human secreted proteins AAB34218-B34264. The genes can be used to  
XX generate fusion proteins by linking to the gene for the human  
XX immunoglobulin G Fc portion (AAC59498) for increasing the stability of  
XX the fusion protein as compared to the human protein only. The genes and  
XX conditions, e.g. by protein or gene therapy. The genes are isolated  
XX from a range of human tissues disclosed in the specification. The  
XX nucleic acids, proteins, antibodies and (ant)agonists are useful in  
XX the diagnosis, treatment and prevention of: (a) cancer, e.g. breast  
XX and ovarian cancer, and other cancers of the adrenal gland, bone  
XX marrow, breast, gastrointestinal tract, liver, lung, or urogenital;  
XX (b) immune disorders e.g. Addison's disease, allergies, autoimmune  
XX haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's  
XX disease, multiple sclerosis, rheumatoid arthritis and ulcerative  
XX colitis; (c) cardiovascular disorders such as myocardial ischaemias;  
XX wound healing; (e) neurological diseases e.g. cerebral anoxia and  
XX epilepsy; and (f) infectious diseases such as viral, bacterial, fungal  
XX and parasitic infections.  
XX  
XX Sequence 785 BP; 173 A; 255 C; 236 G; 121 T; 0 other;

Query Match 86.8%; Score 715.4; DB 21; Length 785;  
Best Local Similarity 99.7%; Pred. No. 9.8e-148;  
Matches 727; Conservative 0; Mismatches 1; Indels 1; Gaps 1;  
QY 96 tcttgccgtgcgcgttcctccggtatccttggtcgtccagcgtggccgcgggtgtgca 155  
Db 23 tcttgccgtgcgcgttcctccggtatccttggtcgtccagcgtggccgcgggtgtgca 82  
QY 156 ggcacgaggtgtccatcagcagcgttgccacagcgtggccacagcagcagcagcagcagcct 215  
Db 83 ggcacgaggtgtccatcagcagcgttgccacagcgtggccacagcagcagcagcagcagcct 142



PR	19-APR-1999;	99US-0130077.
PR	21-APR-1999;	99US-0130049.
PR	23-APR-1999;	99US-0130510.
PR	25-APR-1999;	99US-0130891.
PR	28-APR-1999;	99US-0131149.
PR	30-APR-1999;	99US-0132048.
PR	30-APR-1999;	99US-0132407.
PR	04-MAY-1999;	99US-0132484.
PR	05-MAY-1999;	99US-0132485.
PR	06-MAY-1999;	99US-0132486.
PR	07-MAY-1999;	99US-0132487.
PR	07-MAY-1999;	99US-0132863.
PR	11-MAY-1999;	99US-0134256.
PR	14-MAY-1999;	99US-0134218.
PR	14-MAY-1999;	99US-0134219.
PR	14-MAY-1999;	99US-0134321.
PR	14-MAY-1999;	99US-0134370.
PR	18-MAY-1999;	99US-0134768.
PR	19-MAY-1999;	99US-0134941.
PR	20-MAY-1999;	99US-0135124.
PR	21-MAY-1999;	99US-0135353.
PR	24-MAY-1999;	99US-0135629.
PR	25-MAY-1999;	99US-0136021.
PR	27-MAY-1999;	99US-0136392.
PR	28-MAY-1999;	99US-0136788.
PR	01-JUN-1999;	99US-0137222.
PR	03-JUN-1999;	99US-0137528.
PR	04-JUN-1999;	99US-0137502.
PR	07-JUN-1999;	99US-0137724.
PR	08-JUN-1999;	99US-0138094.
PR	10-JUN-1999;	99US-0138540.
PR	10-JUN-1999;	99US-0138847.
PR	14-JUN-1999;	99US-0139119.
PR	16-JUN-1999;	99US-0139452.
PR	17-JUN-1999;	99US-0139492.
PR	18-JUN-1999;	99US-0139454.
PR	18-JUN-1999;	99US-0139455.
PR	18-JUN-1999;	99US-0139456.
PR	18-JUN-1999;	99US-0139457.
PR	18-JUN-1999;	99US-0139458.
PR	18-JUN-1999;	99US-0139459.
PR	18-JUN-1999;	99US-0139460.
PR	18-JUN-1999;	99US-0139461.
PR	18-JUN-1999;	99US-0139462.
PR	18-JUN-1999;	99US-0139463.
PR	18-JUN-1999;	99US-0139750.
PR	18-JUN-1999;	99US-0139763.
PR	21-JUN-1999;	99US-0139817.
PR	22-JUN-1999;	99US-0139899.
PR	23-JUN-1999;	99US-0140353.
PR	23-JUN-1999;	99US-0140354.
PR	24-JUN-1999;	99US-0140695.
PR	24-JUN-1999;	99US-0140823.
PR	29-JUN-1999;	99US-0140991.
PR	30-JUN-1999;	99US-014287.
PR	01-JUL-1999;	99US-0141842.
PR	01-JUL-1999;	99US-0142154.
PR	02-JUL-1999;	99US-0142055.
PR	06-JUL-1999;	99US-0142390.
PR	08-JUL-1999;	99US-0142803.
PR	09-JUL-1999;	99US-0143277.
PR	12-JUL-1999;	99US-0143542.
PR	13-JUL-1999;	99US-0143624.
PR	14-JUL-1999;	99US-0144005.
PR	15-JUL-1999;	99US-0144086.
PR	16-JUL-1999;	99US-0144087.
PR	19-JUL-1999;	99US-0144325.
PR	19-JUL-1999;	99US-0144331.
PR	19-JUL-1999;	99US-0144332.
PR	19-JUL-1999;	99US-0144333.
PR	19-JUL-1999;	99US-0144334.

PR 13-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 21-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145271.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0145386.  
PR 02-AUG-1999; 99US-0145388.  
PR 03-AUG-1999; 99US-0145389.  
PR 03-AUG-1999; 99US-0147036.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0158293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.

PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 14-OCT-1999; 99US-0159638.  
PR 18-OCT-1999; 99US-0159584.  
PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160768.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 21-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160980.  
PR 22-OCT-1999; 99US-0160981.  
PR 22-OCT-1999; 99US-0160989.  
PR 23-OCT-1999; 99US-0161404.  
PR 23-OCT-1999; 99US-0161405.  
PR 25-OCT-1999; 99US-0161406.  
PR 26-OCT-1999; 99US-0161359.  
PR 26-OCT-1999; 99US-0161360.  
PR 26-OCT-1999; 99US-0161361.  
PR 28-OCT-1999; 99US-0161920.  
PR 28-OCT-1999; 99US-0161992.  
PR 28-OCT-1999; 99US-0161993.  
PR 29-OCT-1999; 99US-0162142.

Query Match 53.8%; Score 602; DB 21; Length 132;

Best Local Similarity 81.8%; Pred. No. 1.7e-59; Mismatches 13; Indels 0; Gaps 0;  
Matches 108; Conservative 11;

QY 82 MTFGLACCAVEMHMAAPRYDMDREGVVFVRASPROSDVMIVAGTLTKMAPALRKVYDQM 141  
Db 1 mtfglaccavemhmtgaarydldrfgilfrpsrqsdcmlvagtltknmapalrkvydqm 60  
QY 142 PEPYVVMGSCANGGGYHYHSYVVRGCDRIVPVDIYIPGCPPTAEALLYGILOLRKI 201  
Db 61 pepyvwismgscangcgyhyhsyvrgrdrivpdyivpgcpptaeallygliqlgkxi 120  
QY 202 KREERLQIWYR 213  
Db 121 nrkdfhlwvnk 132

Search completed: February 12, 2002, 20:37:07  
Job time: 2457 sec





GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 12, 2002, 17:47:55 ; Search time 102.3 Seconds  
(without alignments)  
6905.533 Million cell updates/sec

Title: US-09-525-867-9  
Perfect score: 824  
Sequence: 1 cggtcgagcgctcgagcg.....cgggcaaaaaaaaaaaaaa 824

Scoring table: OLIGO\_NUC  
Gapop 60.0, Gapext 60.0

Searched: 930621 seqs, 428662619 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1861242

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : N\_Geneseq\_1101.\*  
1: /SID22/gcgdata/geneseq/geneseq/NA1980.DAT.\*  
2: /SID22/gcgdata/geneseq/geneseq/NA1981.DAT.\*  
3: /SID22/gcgdata/geneseq/geneseq/NA1982.DAT.\*  
4: /SID22/gcgdata/geneseq/geneseq/NA1983.DAT.\*  
5: /SID22/gcgdata/geneseq/geneseq/NA1984.DAT.\*  
6: /SID22/gcgdata/geneseq/geneseq/NA1985.DAT.\*  
7: /SID22/gcgdata/geneseq/geneseq/NA1986.DAT.\*  
8: /SID22/gcgdata/geneseq/geneseq/NA1987.DAT.\*  
9: /SID22/gcgdata/geneseq/geneseq/NA1988.DAT.\*  
10: /SID22/gcgdata/geneseq/geneseq/NA1989.DAT.\*  
11: /SID22/gcgdata/geneseq/geneseq/NA1990.DAT.\*  
12: /SID22/gcgdata/geneseq/geneseq/NA1991.DAT.\*  
13: /SID22/gcgdata/geneseq/geneseq/NA1992.DAT.\*  
14: /SID22/gcgdata/geneseq/geneseq/NA1993.DAT.\*  
15: /SID22/gcgdata/geneseq/geneseq/NA1994.DAT.\*  
16: /SID22/gcgdata/geneseq/geneseq/NA1995.DAT.\*  
17: /SID22/gcgdata/geneseq/geneseq/NA1996.DAT.\*  
18: /SID22/gcgdata/geneseq/geneseq/NA1997.DAT.\*  
19: /SID22/gcgdata/geneseq/geneseq/NA1998.DAT.\*  
20: /SID22/gcgdata/geneseq/geneseq/NA1999.DAT.\*  
21: /SID22/gcgdata/geneseq/geneseq/NA2000.DAT.\*  
22: /SID22/gcgdata/geneseq/geneseq/NA2001.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Score	Length	ID	Description
1	747	90.7	772	AA158780	Human polynucleoti
2	735	89.2	752	AA160586	Human polynucleoti
3	592	71.8	785	AA160586	Human secreted pro
4	40	4.9	117	AA124101	Human gene signatu
5	27	3.3	789	AA161275	3' CDNA sequence o
6	27	3.3	789	AA158509	3' fragment of pro
7	27	3.3	789	AA1606272	Human immunogenic
8	27	3.3	789	AA10031	Human prostate tum
9	27	3.3	789	AA133388	Human prostate-spe
10	27	3.3	789	AA184702	Human prostate-spe
11	27	3.3	789	AA102453	Prostate tumour an

12	27	3.3	1038	21	AA249556	Partial coding reg
13	27	3.3	1989	20	AA26147	Rat pan-s/tk recep
14	25	3.0	381	21	AA32035	Plant microsatelli
15	24	2.9	1780	21	AA21665	Human breast and o
16	24	2.9	1804	22	AA57575	Human brain cell s
17	23	2.8	1107	21	AA89464	LeuB subunit of 3-
18	23	2.8	2037	22	AA51958	Mycobacterium tube
19	22	2.7	436	21	AA41443	Zea mays DNA fragm
20	22	2.7	597	22	AAH8764	Peppermint plant o
21	22	2.7	631	22	AAH07813	Human cDNA clone (
22	22	2.7	1360	22	AAH42475	cDNA sequence of t
23	22	2.7	1859	22	AA25214	Maize ferulate-5-h
24	22	2.7	1859	22	AA25214	Maize ferulate-5-h
25	22	2.7	1918	22	AAH14977	Human cDNA sequenc
26	22	2.7	2174	22	AAH16635	Human cDNA sequenc
27	22	2.7	2308	15	AAQ66638	Human Activin rece
28	22	2.7	2308	20	AA209845	Human hALK-5 clone
29	22	2.7	2308	21	AA248863	Human Transforming
30	22	2.7	2600	17	AA143624	Chromatin regulato
31	22	2.7	5216	22	AAH28355	Nucleotide sequenc
32	21	2.5	152	21	AA04711	Human secreted pro
33	21	2.5	273	21	AA41424	Zea mays DNA fragm
34	21	2.5	374	21	AA31318	Plant microsatelli
35	21	2.5	379	21	AA31636	Plant microsatelli
36	21	2.5	385	21	AA51712	Zea mays DNA fragm
37	21	2.5	393	22	AA81463	Rice glutaredoxin
38	21	2.5	420	21	AA41346	Zea mays DNA fragm
39	21	2.5	424	21	AA51744	Zea mays DNA fragm
40	21	2.5	428	19	AAV38430	CDNA encoding the
41	21	2.5	456	21	AA39822	Zea mays DNA fragm
42	21	2.5	459	19	AAV38429	CDNA encoding the
43	21	2.5	472	21	AA40956	Zea mays DNA fragm
44	21	2.5	473	21	AA43526	Zea mays DNA fragm
45	21	2.5	482	21	AA43561	Zea mays DNA fragm

ALIGNMENTS

RESULT 1  
AA158780  
ID AA158780 standard; cDNA; 772 BP.  
XX AA158780;  
XX  
XX  
DT 22-OCT-2001 (first entry)  
XX Human polynucleotide SEQ ID NO 983.  
DE  
XX Human; nontropic; immunosuppressant; cytostatic; gene therapy; cancer;  
KW peripheral nervous system; neuropathy; central nervous system; CNS;  
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;  
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;  
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;  
KW leukaemia; ss.  
XX  
OS Homo sapiens.  
XX  
XX WO200153312-A1.  
XX  
XX 26-JUL-2001.  
PD  
XX 26-DEC-2000; 2000WO-US34263.  
PF  
XX 21-JAN-2000; 2000US-0488725.  
PR 25-APR-2000; 2000US-0552317.  
PR 09-JUL-2000; 2000US-0598042.  
PR 19-JUL-2000; 2000US-0620312.  
PR 03-AUG-2000; 2000US-0651450.  
PR 14-SEP-2000; 2000US-0662191.  
PR 19-OCT-2000; 2000US-0693056.  
PR 29-NOV-2000; 2000US-0727344.  
XX

(HYSE-) HYSEQ INC.

PA Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;  
 PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;  
 PI Zhao QA, Zhou P, Goodrich R, Drmanac R;

DR WPI: 2001-442253/47.  
 DR P-PSDB: AAM39624.

XX Novel nucleic acids and polypeptides, useful for treating disorders  
 PT such as central nervous system injuries -

XX Claim 1; SEQ ID NO 983; 10078pp; English.

XX The invention relates to human nucleic acids (AAI57798-AAI61369) and  
 CC the encoded polypeptides (AAM38642-AAM42213) with neurotropic,  
 CC immunosuppressant and cytostatic activity. The polynucleotides are useful  
 CC in gene therapy. A composition containing a polypeptide or polynucleotide  
 CC of the invention may be used to treat diseases of the peripheral nervous  
 CC system, such as peripheral nervous injuries, peripheral neuropathy and  
 CC localised neuropathies and central nervous system diseases, such as  
 CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic  
 CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the  
 CC utilisation of the activities such as: Immune system suppression,  
 CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic  
 CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,  
 CC assays for receptor activity, arthritis and inflammation, leukaemias and  
 CC C.N.S disorders.

CC Note: The sequence data for this patent did not form part of the printed  
 CC specification.

XX Sequence 772 BP; 136 A; 265 C; 245 G; 126 T; 0 other;

Query Match 90.7%; Score 747; DB 22; Length 772;  
 Best Local Similarity 100.0%; Pred. No. 5,6e-290;  
 Matches 747; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 53 gaagggcagccaaagatggcgtgtgtcagctctgtgctgcggttcggtatct 122  
 DB 8 gaagggcagccaaagatggcgtgtgtcagctctgtgctgcggttcggtatct 67  
 QY 123 tggctcgtcgtccagcgtggcgcgtgtgcagggcagcaggtgtccatcagagcgtgc 182  
 DB 58 tggctcgtcgtccagcgtggcgcgtgtgcagggcagcaggtgtccatcagagcgtgc 127  
 QY 183 caccgatggcccaagcagcaccagcgtgcctgccaaagcgcagcgcgtggtcccaaa 242  
 DB 128 caccgatggcccaagcagcaccagcgtgcctgccaaagcgcagcgcgtggtcccaaa 187  
 QY 243 acccagcagcggggcaggtatgtgtggtccaaagcgtggatgacctgctcaactgggccc 302  
 DB 188 acccagcagcggggcaggtatgtgtggtccaaagcgtggatgacctgctcaactgggccc 247  
 QY 303 ccggagttctctgtggcccatgacctcgcgtgcgtgcgtgcgtgcgtgcgtgcgtgcgtgc 362  
 DB 248 ccggagttctctgtggcccatgacctcgcgtgcgtgcgtgcgtgcgtgcgtgcgtgcgtgc 307  
 QY 363 catggcagcaccgcgtacagacatggacgcgttcttggcgtgttcttcgcgcgcgcgcgcgc 422  
 DB 308 catggcagcaccgcgtacagacatggacgcgttcttggcgtgttcttcgcgcgcgcgcgcgc 367  
 QY 423 ccagtcgcagcgtcatgctgtgc 482  
 DB 368 ccagtcgcagcgtcatgctgtgc 427  
 QY 483 caaggtctacgacagatgc 542  
 DB 428 caaggtctacgacagatgc 487  
 QY 543 cggaggaggtactacacattctcactcgtgtgtgtggggcgtgcgcgcgcgcgcgcgcgcgcgc 602  
 DB 488 cggaggaggtactacacattctcactcgtgtgtgtggggcgtgcgcgcgcgcgcgcgcgcgcgc 547

QY 603 cgtggacatctacatccagcgtgccacactacgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 662  
 DB 548 cgtggacatctacatccagcgtgccacactacgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 607  
 QY 663 gcagctgcagaggaagatcaagcggagcggagcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 722  
 DB 608 gcagctgcagaggaagatcaagcggagcggagcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 667  
 QY 723 ccgc 782  
 DB 668 ccgc 727  
 QY 783 aggtgtcaataaactgc 809  
 DB 728 aggtgtcaataaactgc 754

## RESULT 2

AAI60566

ID AAI60566 standard; cDNA; 752 BP.

XX AAI60566;

AC AAI60566;

XX 22-OCT-2001 (first entry)

XX Human polynucleotide SEQ ID NO 4555.

XX Human; neurotropic; immunosuppressant; cytostatic; gene therapy; cancer;  
 KW peripheral nervous system; neuropathy; central nervous system; CNS;  
 KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;  
 KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;  
 KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;  
 KW leukaemia; ss.

XX Homo sapiens.

XX WO200153312-A1.

XX 26-JUL-2001.

XX 26-DEC-2000; 2000WO-US34263.

XX 21-JAN-2000; 2000US-0488725.

XX 25-APR-2000; 2000US-0552317.

XX 09-JUL-2000; 2000US-0598042.

XX 19-JUL-2000; 2000US-0620312.

XX 03-AUG-2000; 2000US-0653450.

XX 14-SEP-2000; 2000US-0662191.

XX 19-OCT-2000; 2000US-0693036.

XX 29-NOV-2000; 2000US-0727344.

XX (HYSE-) HYSEQ INC.

XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;

XX Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;

XX Zhao QA, Zhou P, Goodrich R, Drmanac R;

XX WPI: 2001-442253/47.

XX P-PSDB: AAM41410.

XX Novel nucleic acids and polypeptides, useful for treating disorders

XX such as central nervous system injuries -

XX Claim 1; SEQ ID NO 4555; 10078pp; English.

XX The invention relates to human nucleic acids (AAI57798-AAI61369) and  
 CC the encoded polypeptides (AAM38642-AAM42213) with neurotropic,  
 CC immunosuppressant and cytostatic activity. The polynucleotides are useful  
 CC in gene therapy. A composition containing a polypeptide or polynucleotide  
 CC of the invention may be used to treat diseases of the peripheral nervous  
 CC system, such as peripheral nervous injuries, peripheral neuropathy and  
 CC localised neuropathies and central nervous system diseases, such as

```
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC utilisation of the activities such as: immune system suppression,
CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC assays for receptor activity, arthritis and inflammation, leukaemias and
CC C.N.S disorders.
CC Note: The sequence data for this patent did not form part of the printed
CC specification.
XX Sequence 752 BP; 130 A; 260 C; 236 G; 126 T; 0 other;
SQ

Query Match      89.2%; Score 735; DB 22; Length 752;
Best Local Similarity 100.0%; Pred. No. 3.5e-285;
Matches 735; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 75 caagatggcgggtgctgcagctccctggcgtgcggcgttcctcggtacccctgtgtcgtcgctc 134
   |||
Db 1 caagatggcgggtgctgcagctccctggcgtgcggcgttcctcggtacccctgtgtcgtcgctc 60

QY 135 cagctggcgcgcgtgtcagcagcagaggtgtccatcagagcgtggccacccgatggccc 194
   |||
Db 61 cagctggcgcgcgtgtcagcagcagaggtgtccatcagagcgtggccacccgatggccc 120

QY 195 aagcagcaccagcctgcctcccaagccagagccgtggctcccaaccacccagcagcgg 254
   |||
Db 121 aagcagcaccagcctgcctcccaagccagagccgtggctcccaaccacccagcagcgg 180

QY 255 gggcagatgtggtggccaaagtggatgaacctgtcaactgggcccgcggagttctct 314
   |||
Db 181 gggcagatgtggtggccaaagtggatgaacctgtcaactgggcccgcggagttctct 240

QY 315 gtggcccatgaacttggcctgctgcctgcctggcctggagatgatgcacatggcagcacc 374
   |||
Db 241 gtggcccatgaacttggcctgctgcctggcctggagatgatgcacatggcagcacc 300

QY 375 ccgtacagacatggaccgctttggcgtgttcttcgcgcgccagccgcgcagtcgcagct 434
   |||
Db 301 ccgtacagacatggaccgctttggcgtgttcttcgcgcgccagccgcgcagtcgcagct 360

QY 435 catgctgctggcgcgcacactcaccacaagatggccacagcgttcgcaaggtctacga 494
   |||
Db 361 catgctgctggcgcgcacactcaccacaagatggccacagcgttcgcaaggtctacga 420

QY 495 ccagatgcggagcgcgcgtactgtgtctccatggggagctgcgccaaagagagagcta 554
   |||
Db 421 ccagatgcggagcgcgcgtactgtgtctccatggggagctgcgccaaagagagagcta 480

QY 555 ctaccactattctcgtgtgtgaggggtgcacgcagctggtcccggtggacatcta 614
   |||
Db 481 ctaccactattctcgtgtgtgaggggtgcacgcagctggtcccggtggacatcta 540

QY 615 catccagagctgccacactacggccagcgcctgtctctacggcatcctgcagctgcagag 674
   |||
Db 541 catccagagctgccacactacggccagcgcctgtctctacggcatcctgcagctgcagag 600

QY 675 gaagatcaagcgggagcggagcgtcagatgtgtacgcagggtagcgcgcgccccgcgg 734
   |||
Db 601 gaagatcaagcgggagcggagcgtcagatgtgtacgcagggtagcgcgcgccccgcgg 660

QY 735 ccgcggagcctgtgcgcctctctgtccacagcctgtgttcccgtagagtttcaata 794
   |||
Db 661 ccgcggagcctgtgcgcctctctgtccacagcctgtgttcccgtagagtttcaata 720

QY 795 aacctgcccctcgggc 809
   |||
Db 721 aacctgcccctcgggc 735

RESULT 3
AAC59553
ID AAC59553 standard; cDNA; 785 BP.
```

```
XX AAC59553;
AC 15-FEB-2001 (first entry)
DT Human secreted protein cDNA sequence #47.
DE
XX Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
XX antiallergic; hepatotropic; antidiabetic; antiinflammatory; antiulcer;
XX vulnary; anticonvulsant; antibacterial; antifungal; antiparasitic;
XX cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;
XX neurological disease; infection; human; secreted protein; ss.
XX Homo sapiens.
OS
XX WO200055352-A2.
PN
XX 21-SEP-2000.
PD
XX 09-MAR-2000; 2000WO-US06044.
PF
XX 12-MAR-1999; 99US-0124099.
PR 03-DEC-1999; 99US-0168661.
XX (HUMA-) HUMAN GENOME SCI INC.
PA
XX Rosen CA, Ruben SM, Komatsoulis G;
PI WPI; 2000-502124/57.
XX P-PSDB; AAB34261.
DR
XX Novel human secreted proteins useful for diagnosis, prevention and
XX treatment of disorders including neurological, cell proliferative,
XX cardiovascular, autoimmune and inflammatory disorders and microbial
XX infections
XX Claim 1: Page 345; 383pp; English.
XX The invention relates to the isolation of genes AAC59507-C59556 encoding
XX 50 human secreted proteins AAB34218-A34264. The genes can be used to
XX generate fusion proteins by linking to the gene for the human
XX immunoglobulin G Fc portion (AAC59498) for increasing the stability of
XX the fusion protein as compared to the human protein only. The genes and
XX proteins are useful for preventing, ameliorating or treating medical
XX conditions, e.g. by protein or gene therapy. The genes are isolated
XX from a range of human tissues disclosed in the specification. The
XX nucleic acids, proteins, antibodies and (ant)agonists are useful in
XX the diagnosis, treatment and prevention of: (a) cancer, e.g. breast
XX and ovarian cancer, and other cancers of the adrenal gland, bone, bone
XX marrow, breast, gastrointestinal tract, liver, lung, or urogenital;
XX (b) immune disorders e.g. Addison's disease, allergies, autoimmune
XX haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's
XX disease, multiple sclerosis, rheumatoid arthritis and ulcerative
XX colitis; (c) cardiovascular disorders such as myocardial ischaemias; (d)
XX wound healing; (e) neurological diseases e.g. cerebral anoxia and
XX epilepsy; and (f) infectious diseases such as viral, bacterial, fungal
XX and parasitic infections.
XX Sequence 785 BP; 173 A; 255 C; 236 G; 121 T; 0 other;
SQ
```

```
Query Match      71.8%; Score 592; DB 21; Length 785;
Best Local Similarity 99.9%; Pred. No. 5.9e-228;
Matches 712; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 96 tcttgccctgcgcggttcctcggtacccctgtgtcgtcgtccagcgtggcccggtgtgca 155
   |||
Db 23 tcttgccctgcgcggttcctcggtacccctgtgtcgtcgtccagcgtggcccggtgtgca 82

QY 156 ggcacgaggtgtccatcagcgtgtggccacgcagtgcccaagcagccacccagcctgcct 215
   |||
Db 83 ggcacgaggtgtccatcagcgtgtggccacgcagtgcccaagcagccacccagcctgcct 142
```

QY	216	gcaaaagccacgagccgttggtctcccaaacccacgacgcgggcgagttatgtgtggtgccaa	275
Db	143	gcaaaagccacgagccgttggtctcccaaacccacgacgcgggcgagttatgtgtggtgccaa	202
QY	276	gctgaatgaacctgtcaactggccgcgcgcgaggttctctgttgcccatgaccttgcgct	335
Db	203	gctggtatgaacctgtcaactggccgcgcgcgaggttctctgttgcccatgaccttgcgct	262
QY	336	ggcctgtgcgccttgagatgatcacatggcagcaccccgctaccacatgacacgctt	395
Db	263	ggcctgtgcgccttgagatgatcacatggcagcaccccgctaccacatgacacgctt	322
QY	396	tggcgttggtcttcgcac	455
Db	323	tggcgttggtcttcgcac	382
QY	456	caccacaagaatggccccacgcgttcgcgaaggtctacgacacagatgcgcgagcgcgcta	515
Db	383	caccacaagaatggccccacgcgttcgcgaaggtctacgacacagatgcgcgagcgcgcta	442
QY	516	cgttggtctccatggggagcttgcccaacggagaggtactaccactattcctactcgt	575
Db	443	cgttggtctccatggggagcttgcccaacggagaggtactaccactattcctactcgt	502
QY	576	ggtgagggggtgcgacgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcac	635
Db	503	ggtgagggggtgcgacgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcac	561
QY	636	ggcccagggcctgtctacggcatcctgcagctgcagaggagatcaaggcgagcgagag	695
Db	562	ggcccagggcctgtctacggcatcctgcagctgcagaggagatcaaggcgagcgagag	621
QY	696	gctgcagatcttgtaaccgcaggttagcgccgcgcgcgcgcgcgcgcgcgcgcgcgc	755
Db	622	gctgcagatcttgtaaccgcaggttagcgccgcgcgcgcgcgcgcgcgcgcgcgcgc	681
QY	756	ctgtccccagcctgtgtgtcccgtaggttgtaataaacctgcgcctcggg	808
Db	682	ctgtccccagcctgtgtgtcccgtaggttgtaataaacctgcgcctcggg	734

## RESULT

RESOL<sup>4</sup>  
AAT24101  
ID AAT24101 standard; cDNA to mRNA; 117 BP.

DT 06-SEP-1996 (first entry)

Gene signature; messenger RNA; mRNA; relative abundance; frequency;  
human; cloning; mapping; non-biased library; diagnosis; detection;  
cell typing; abnormal cell function; ss.

XX  
OS  
Homo sapiens.

XX PN WO9514772-A1.

XX  
PD 01-JUN-1995.

XX  
PF 11-NOV-1994: 94WO-JP01916.

XX  
PR 12-NOV-1993: 93JP-0355504.

XX PA (MATS/) MATSUBARA K.

XX  
РА (ОКУВ/) ОКУВО К.

PI Matsubara K, Okubo K;  
YX

DR WPI; 1995-206931/27.  
XX

PT Identifying gene signatures in 3'-directed human cDNA library - e.g.

for diagnosis of abnormal cell function, by preparing cDNA that reflects relative abundance of corresp. mRNA in specific human tissues

Claim 1; Page 1530; 2245pp; Japanese.

A single-stranded DNA (or its complementary strand or the corresp. double-stranded DNA) which comprises one of the 7837 "GS" sequences given in AAT19001-426937 and which is able to hybridise to part of human genomic DNA, cDNA or mRNA is claimed. The GS (Gene Signature) sequences were obtained from 3'-directed cDNA libraries prepared from various human tissues; synthesis of cDNA was initiated from the 3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-untranslated sequence is unique to a particular mRNA species, almost all the 3'-oriented cDNAs hybridise with specific mRNAs. Each library is constructed so as to reflect accurately the relative abundance of different mRNAs in the particular tissue from which it was derived. The appearance frequency of a given GS in a cDNA library can be determined (esp. using primers and probes derived from the GS sequences) as a means of diagnosing abnormal cell function or for recognising different cell types.

Sequence 117 BP; 15 A; 43 C; 34 G; 23 T; 2 other;

## RESULT 5

AAV61275/C

ID AAV61275 standard; cDNA; 789 BP.

AA  
AC  
AAV61275;

DT 06-JAN-1999 (first entry)

3' cDNA sequence of prostate tumour clone J1-21.

Prostate: cancer: tumour: vaccine: immunogen: clone: ss:  
XX  
KW

XX  
OS Homo sapiens.

XX PN WO9837093-A2.

XX  
PD 27-AUG-1998XX  
PF  
25-FEB-1998: 98W0-IIS03492XX  
09-FEB-1998. 08US-0020956

PR 25-FEB-1997; 97US-0806099.  
PR 01-AUG-1997; 97US-0904804

XX  
DA (CORT-) CORTVA CORR

XX	DT	Y:	Z:
DT	DT	DT	DT

XX  
DD  
WDT: 1008-600806 /

XX  
DE

PT used in a vaccine for the treatment of prostate cancer  
XX  
PS Claim 12; Page 48; 130pp; English.

The present sequence is a DNA which encodes an immunogenic portion of a prostate tumour protein. The encoded immunogen, or the DNA itself, can be used as a vaccine for the treatment of prostate cancer. The DNA was identified by analysis of a subtracted cDNA library obtained by



XX WI; 2001-425873/45.  
XX  
XX  
PT New polynucleotide encoding a prostate-specific protein, for  
PT diagnosing, monitoring and treating prostate cancer in a patient and  
PT for use in vaccines -  
XX  
XX Claim 1; Page 240; 543pp; English.  
PS  
PS  
XX The present invention describes polynucleotide sequences (I) which encode  
CC prostate-specific proteins (II). (I) and (II) have cytostatic activity,  
CC and can be used in vaccine production and gene therapy. (I), (II),  
CC antibodies to (II), fusion proteins comprising (II), and isolated  
CC T cells prepared using (I) or (II) are used treat cancer in a patient.  
CC (I) and the antibodies are also used in the detection of cancer in a  
CC patient. The cancer that is diagnosed or treated is particularly  
CC prostate cancer. (I) and (II) can be used in vaccines. The antibodies or  
CC (I) can be used for monitoring the progression of cancer in a patient.  
CC (I) and (II) can also be used to improve diagnostic and therapeutic  
CC methods for prostate cancer. They can indicate the level of metastasis  
CC as well as the prostate volume. AAH93357 to AAH93944 and AAM01115 to  
CC AAM01318 represent polynucleotide and amino acid sequences used in the  
CC exemplification of the present invention.

XX Sequence 789 BP; 129 A; 240 C; 159 G; 195 T; 66 other;  
XX

RESULT	10	
AAH84702/C		
ID	AAH84702 standard; cDNA; 789 BP.	
XX		
XX		
AC	AAH84702;	
XX		
XX		
DT	25-SEP-2001 (first entry)	
XX		
XX		
DE	Human prostate-specific 3' cDNA sequence J1-21.	
XX		
XX		
KW	Human; prostate cancer; therapy; diagnosis; cat eye syndrome;	
KW	chromosome 22q11.2; prostate-specific protein; chromosome 1;	
KW	prostate specific antigen; PSA; ss.	

DE	Human prostate-specific 3' CDNA sequence JI-21.
XX	
XX	Human; prostate cancer; therapy; diagnosis; cat eye syndrome;
KW	chromosome 22q11.2; prostate-specific protein; chromosome 1;
KW	prostate specific antigen; PSA; ss.
XX	
XX	Homo sapiens.
OS	
XX	WO200134802-A2.
PN	
XX	17-MAY-2001.
XX	
PD	09-NOV-2000; 2000WO-US30904.
XX	
XX	12-NOV-1999; 99US-0439313.
PR	18-NOV-1999; 99US-0443556.
PR	
XX	(CORI-) CORIXA CORP.
XX	
XX	Xu J, Dillon DC, Mitcham JL, Harlocker SL, Jiang Y, Reed SG;
PI	Kalos MD, Retter MW, Stolk JA, Day CH, Skeiky YAW, Wang A,
PI	
XX	WPI; 2001-308785/32.
XX	
XX	Isolated polypeptide comprising at least an immunogenic portion of a
PT	prostate-specific protein, useful in the diagnosis and therapy of
PT	prostate cancer -
XX	
XX	Claim 31; Page 140; 325pp; English.
PS	

```

XX
SQ Sequence 789 BP; 129 A; 240 C; 159 G; 195 T; 66 other;

Query Match          3.3%; Score 27; DB 22; Length 789;
Best Local Similarity 100.0%; Pred. NO. 0.071;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 722 gccgcgcgcgcgcgcgcgcgcgcctgt 748
      ||| ||| ||| ||| ||| ||| |||
Db 151 GCCGCGCGCGCGCGCGGAGCCTGT 125

```

RESULT 12  
AAZ49556  
ID AAZ49556 standard: DNA: 1038 BP.

XX	AA249556;
AC	
XX	30-MAR-2000
DT	(first entry)

XX	Partial coding region of human latent
DE	
XX	Fragmentation vector: target sequence
KW	

KW telomere; homologous recombination; deletion fragment; contig mapping;  
KW positional information; CCG repeat; human; Transforming growth factor;  
KW Fragmented YAC set; pDVCCG vector; latent TGF-beta binding protein; ds.  
XX

Homo sapiens.

Key	Location/Qualifiers
repeat_region	301..321

```
repeat unit
/*tag= a
/rpt_type= TANDEM
/note="7 CCG triplet repeat"
301..303
```

```
repeat_unit
note="/ CCG triplet repeat"
301..303
/*tag= b
```

3  
 4  
 5  
 6  
 7  
 8  
 9  
 10  
 11  
 12  
 13  
 14  
 15  
 16  
 17  
 18  
 19  
 20  
 21  
 22  
 23  
 24  
 25  
 26  
 27  
 28  
 29  
 30  
 31  
 32  
 33  
 34  
 35  
 36  
 37  
 38  
 39  
 40  
 41  
 42  
 43  
 44  
 45  
 46  
 47  
 48  
 49  
 50  
 51  
 52  
 53  
 54  
 55  
 56  
 57  
 58  
 59  
 60  
 61  
 62  
 63  
 64  
 65  
 66  
 67  
 68  
 69  
 70  
 71  
 72  
 73  
 74  
 75  
 76  
 77  
 78  
 79  
 80  
 81  
 82  
 83  
 84  
 85  
 86  
 87  
 88  
 89  
 90  
 91  
 92  
 93  
 94  
 95  
 96  
 97  
 98  
 99  
 100  
 101  
 102  
 103  
 104  
 105  
 106  
 107  
 108  
 109  
 110  
 111  
 112  
 113  
 114  
 115  
 116  
 117  
 118  
 119  
 120  
 121  
 122  
 123  
 124  
 125  
 126  
 127  
 128  
 129  
 130  
 131  
 132  
 133  
 134  
 135  
 136  
 137  
 138  
 139  
 140  
 141  
 142  
 143  
 144  
 145  
 146  
 147  
 148  
 149  
 150  
 151  
 152  
 153  
 154  
 155  
 156  
 157  
 158  
 159  
 160  
 161  
 162  
 163  
 164  
 165  
 166  
 167  
 168  
 169  
 170  
 171  
 172  
 173  
 174  
 175  
 176  
 177  
 178  
 179  
 180  
 181  
 182  
 183  
 184  
 185  
 186  
 187  
 188  
 189  
 190  
 191  
 192  
 193  
 194  
 195  
 196  
 197  
 198  
 199  
 200  
 201  
 202  
 203  
 204  
 205  
 206  
 207  
 208  
 209  
 210  
 211  
 212  
 213  
 214  
 215  
 216  
 217  
 218  
 219  
 220  
 221  
 222  
 223  
 224  
 225  
 226  
 227  
 228  
 229  
 230  
 231  
 232  
 233  
 234  
 235  
 236  
 237  
 238  
 239  
 240  
 241  
 242  
 243  
 244  
 245  
 246  
 247  
 248  
 249  
 250  
 251  
 252  
 253  
 254  
 255  
 256  
 257  
 258  
 259  
 260  
 261  
 262  
 263  
 264  
 265  
 266  
 267  
 268  
 269  
 270  
 271  
 272  
 273  
 274  
 275  
 276  
 277  
 278  
 279  
 280  
 281  
 282  
 283  
 284  
 285  
 286  
 287  
 288  
 289  
 290  
 291  
 292  
 293  
 294  
 295  
 296  
 297  
 298  
 299  
 300  
 301  
 302  
 303  
 304  
 305  
 306  
 307  
 308  
 309  
 310  
 311  
 312  
 313  
 314  
 315  
 316  
 317  
 318  
 319  
 320  
 321  
 322  
 323  
 324  
 325  
 326  
 327  
 328  
 329  
 330  
 331  
 332  
 333  
 334  
 335  
 336  
 337  
 338  
 339  
 340  
 341  
 342  
 343  
 344  
 345  
 346  
 347  
 348  
 349  
 350  
 351  
 352  
 353  
 354  
 355  
 356  
 357  
 358  
 359  
 360  
 361  
 362  
 363  
 364  
 365  
 366  
 367  
 368  
 369  
 370  
 371  
 372  
 373  
 374  
 375  
 376  
 377  
 378  
 379  
 380  
 381  
 382  
 383  
 384  
 385  
 386  
 387  
 388  
 389  
 390  
 391  
 392  
 393  
 394  
 395  
 396  
 397  
 398  
 399  
 400  
 401  
 402  
 403  
 404  
 405  
 406  
 407  
 408  
 409  
 410  
 411  
 412  
 413  
 414  
 415  
 416  
 417  
 418  
 419  
 420  
 421  
 422  
 423  
 424  
 425  
 426  
 427  
 428  
 429  
 430  
 431  
 432  
 433  
 434  
 435  
 436  
 437  
 438  
 439  
 440  
 441  
 442  
 443  
 444  
 445  
 446  
 447  
 448  
 449  
 450  
 451  
 452  
 453  
 454  
 455  
 456  
 457  
 458  
 459  
 460  
 461  
 462  
 463  
 464  
 465  
 466  
 467  
 468  
 469  
 470  
 471  
 472  
 473  
 474  
 475  
 476  
 477  
 478  
 479  
 480  
 481  
 482  
 483  
 484  
 485  
 486  
 487  
 488  
 489  
 490  
 491  
 492  
 493  
 494  
 495  
 496  
 497  
 498  
 499  
 500  
 501  
 502  
 503  
 504  
 505  
 506  
 507  
 508  
 509  
 510  
 511  
 512  
 513  
 514  
 515  
 516  
 517  
 518  
 519  
 520  
 521  
 522  
 523  
 524  
 525  
 526  
 527

```
/*tag= c
/notes= "Sequence obtained
```

## Fragmente

WO9966059-A1.

23-DEC-1999. 11-JUN-1999: 99WO-EP04106.

II-JUN-1999; 99WO-EP04106.  
12-JUN-1998; 98EP-0201976.

12-JUN-1998; 98EP-0201976.  
(VLA-) VLAAMS INTERUNIVERSITAIR  
Del Favero J, Van Broeckhoven C;

(VLAAR-) VLAAMS INTERUNIVERSITAIR INST BIOTECHNOLOG

Fragmentation vector for producing deletion fragments of yeast artificial chromosomes for genome mapping -  
 Del-Favero J, van Broeckhoven C;  
 WPI: 2000-116773/10.

Fragmentation vector for producing artificial chromosomes for gene transfer

artificial chromosomes for genome mapping -  
Example; Fig 7; 52pp; English.

The present DNA sequence is the partial coding region of human latent transforming 'Growth factor (TGF) beta binding protein, obtained by fragmentation of YAC clones, using the vector pDVCCG. The CCG repeats could be isolated from equally sized fragmented YACs. The regions of recombination contain one or a few CCG/CCG repeats, but are located in the GC rich region. The fragmentation vectors can be used for producing deletion fragments of YACs. The vectors can be used for the sequencing of at least a part of the human genomic sequences present in the YACs. New markers from the chromosomes of the human genome can be isolated and positional information of isolated sequences can be provided, for use in





Search completed: February 12, 2002, 19:56:04  
Job time: 7689 sec



141 gggccggtgtgcagcacgagggtgtccatcagagcgctggccaccgatggcccgaagcag 200  
||| ||| | ||| | ||| | ||| | ||| | ||| |  
739 GGGGATGGTGNNAAGANGNNCTTTTCNTTAGAGNNGGGCCACCGGGGGGCCCAAGGG 680



[illegible]

```

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/587,179
FILING DATE: 20-SEP-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 921,730
FILING DATE: 20-OCT-1986
APPLICATION NUMBER: 597,784
FILING DATE: 06-APR-1984
SEQ ID NO: 2
LENGTH: 3472
5244792-2

```

Query Match 5.6%; Score 46; DB 6; Length 3472;  
Best Local Similarity 43.9%; Pred. No. 0.049;  
Matches 196; Conservative 0; Mismatches 250; Indels 0; Gaps 0;

QY 247 agcagcggggcagatgtgtgtggcccaagctggatgacatcgtcaactggcccgccgg 306  
Db 1116 aactgcattcagaggggtggagcgcgcgtcgtgtacccctacagagattgtgtg 1175  
QY 307 agttctctgtggccatgacattcgtccctggcctgtgtgcgcgtggagatgatgacatg 366  
Db 1176 gcaactggcagactttgtatattcccgctttacggctacagggaggggtcgcacacc 1235  
QY 367 gcagaccccgctacagatcggaccgtttgtgtgtgttcccgccagcccgccag 426  
Db 1236 gaacacacagctacaccccgccgacccgttcaagcaggtgacgggttctacgcgcgac 1295  
QY 427 tccgaactcagatcgtggcggcgcacactcaccaacaaagatggcccgcttcgcaag 486  
Db 1296 ctcaaccaaaggccggccacggcgccgacaccccggaacctgctcagcaccaccaag 1355  
QY 487 gtctacacacagatgcggagccgcgcgtcagttgttccatggggagctggcccaacgga 546  
Db 1356 ttcacgtgctggagactgggtgcacaaagcccgctcgtgtgcacctgacccaagtgg 1415  
QY 547 ggaggtactaccactattcctactcgtgtgtgtgagggctgcgacgcacatgcccgtg 606  
Db 1416 caggagtgacagagatgctgcgtcgcagtagcggcgtctcctcgtatctcctcccgac 1475  
QY 607 gacatctacatccagcgtgcccactacgacggcggcgcctcgtctcagggcatcctgag 666  
Db 1476 gccatctccacccttcac 1535  
QY 667 ctgcagaggagatcaagcggagggc 692  
Db 1536 ggggactgcacggcgaaggacggcccg 1561

RESULT 6  
US-09-593-589-3  
; Sequence 3, Application US/09593589  
; Patent No. 6306655  
; GENERAL INFORMATION:  
; APPLICANT: Brett P. Monia  
; APPLICANT: Madeline M. Butler  
; APPLICANT: Jacqueline Wyatt  
; TITLE OF INVENTION: ANTISENSE MODULATION OF C/EBP ALPHA EXPRESSION  
; FILE REFERENCE: RTS-0119  
; CURRENT APPLICATION NUMBER: US/09/593,589  
; CURRENT FILING DATE: 2000-06-13  
; NUMBER OF SEQ ID NOS: 94  
; SEQ ID NO 3  
; LENGTH: 3318  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (592)...(1658)  
US-09-593-589-3

Query Match 5.5%; Score 45.4; DB 4; Length 3318;  
Best Local Similarity 46.2%; Pred. No. 0.067;  
Matches 188; Conservative 0; Mismatches 216; Indels 3; Gaps 1;

QY 326 ccttcggcctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctg 385  
Db 680 ccttcggccttcccgggggcggggcggcggcggcggcggcggcggcggcggcggcggcgg 739  
QY 386 tggacgcctttgctggtgttcttcggcgcagcccgccagctccgacatcgtatgctgg 445  
Db 740 agccgctggcggcactcgtcggagacagacacgtccatcgtacatcgtacatcgtacac 799

QY 446 cgggcacactcaccaacaagatggccccagcgtcttcgcaaggtctacgacacagatgcggg 505  
Db 800 cggcgccttcaacgacagagttcttcggccgacactgttcagacacagcggcgacagagaga 859  
QY 506 agcc---gcgctcgtggtctccatgggagctggcccaacggagggaggtactactaccact 562  
Db 860 agcccaaggcggcgtgggccccacggggcgggcgggcgggcgactttgactaccgg 919  
QY 563 attcctactcgtggtgaggggctgcgacgcgcatcgtgcctgtgacatctacatccccag 622  
Db 920 gcggccccggggccccggcgcgccgtcatgccggggggagcgacgggcccccgcccg 979  
QY 623 gctgccccactacggcggcggccctgtctctacggcctcgtcagctgcagaggagagatca 682  
Db 980 gctacggcgtgcggcgccggcgtacactggacggcaggtgtgagccctgtacgagcgcg 1039  
QY 683 agcgggagggagggctgcagatctgtgtacgcaggtagcgcgcgcg 729  
Db 1040 tcggggcgccggcgtcgtgcggccgctgtgtatcaagcagggagccccgcg 1086

RESULT 7  
US-08-804-439A-13  
; Sequence 13, Application US/08804439A  
; Patent No. 6015565  
; GENERAL INFORMATION:  
; APPLICANT: Rose, Timothy M.  
; APPLICANT: Bosch, Marnix L.  
; APPLICANT: Strand, Kurt  
; TITLE OF INVENTION: GLYCOPROTEIN B OF THE RPHV/KSHV  
; TITLE OF INVENTION: SUBFAMILY OF HERPES VIRUSES  
; NUMBER OF SEQUENCES: 113  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fish & Richardson P.C.  
; STREET: 4225 Executive Square, Ste 1400  
; CITY: La Jolla  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 92037  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/804,439A  
; FILING DATE: February 21, 1997  
; CLASSIFICATION: 424  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Haile, Lisa A.  
; REGISTRATION NUMBER: 38,347  
; REFERENCE/DOCKET NUMBER: 09176/004001  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (619) 678-5070  
; TELEFAX: (619) 678-5099  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 13:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 2713 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
US-08-804-439A-13

Query Match 5.4%; Score 44.4; DB 3; Length 2713;  
Best Local Similarity 43.7%; Pred. No. 0.11;  
Matches 195; Conservative 0; Mismatches 251; Indels 0; Gaps 0;

QY 247 agcagcggggcagatgtgtgtggcccaagctggatgacatcgtcactggcccgccgg 306  
Db 805 AACTGCATCGTCGAGAGGTGACGCGCGCTCGGTGTACCCGTACGACGAGTTGTGTGCTG 864













Search completed: F



```

|||||
111 TGGCCACCGATGGCCCAAGCAGCACCAGCGCTGCCCTGCCAAGGCCAGA 160
|||||
51 AlaValAlaProLysProSerArgGlyGluTyrValValAlaLysLe 67
|||||
161 GCCGTGGCTCCCAACCCAGCAGCGCGGCGAGTAGTGTGGTGCACAGCT 210
|||||
67 uAspAspLeuValAsnTrpAlaArgArgSerSerLeuTrpProMetThrP 84
|||||
211 GGATGACCTGCTCAACTGGGCGCGCGGAGTCTCTGTGGCCCATGACCT 260
|||||
84 heGlyLeuAlaCysCysAlaValGluMetMetHisMetAlaAlaProArg 100
|||||
261 TCGGCCCTGGCTGTGGCGCGTGGAGATGATGCATGGCAGCACCCCGC 310
|||||
101 TyrAspMetAspArgPheGlyValValPheArgAlaSerProArgGlnSe 117
|||||
311 TACGACATGAGCCGCTTGTGGCTGTCTCCGCGCAGCGCCGCCAGTC 360
|||||
117 rAspValMetIleValAlaGlyThrLeuThrAsnLysMetAlaProAla 134
|||||
361 CGAGCTCATGATCGTGGCGGCACACATCAACACAGATGGCCCCAGCGC 410
|||||
134 euArgLysValTyrAspGlnMetProGluProArgTyrValValSerMet 150
|||||
411 TTCGAAGGCTACGACACAGATCGCGAGCGCGCTACGTGGTCTCCATG 460
|||||
151 GlySerCysAlaAsnGlyGlyTyrTyrHisTyrSerTyrSerValVa 167
|||||
461 GGGAGCTGCGCAACAGGAGGAGGCTACTACCATATTCCTACTCGTGTG 510
|||||
167 lArgGlyCysAspArgIleValProValAspIleTyrIleProGlyCysP 184
|||||
511 GAGGGCTGGACCGCATCGTGGCCGTGGACATCTACATCCAGGCTGCC 560
|||||
184 rOPrThrAlaGluAlaLeuLeuTyrGlyIleLeuGlnLeuGlnArgLys 200
|||||
561 CACCTACGGCGGAGGCGCTCTCTACGGCATCTCGAGCTGCAGAGGAAG 610
|||||
201 lIleLysArgGluArgArgLeuGlnIleTrpTyrArgArg 213
|||||
611 ATCAAGCGGAGCGGAGGCTGCAGATCTGTACCGCAGG 649

```

seq\_name: gb\_est2:BG334974

```

seq_documentation_block:
LOCUS      BG334974          704 bp      mRNA          EST          27-FEB-2001
DEFINITION 602403406F1 NIH_MGC_21 Homo sapiens cDNA clone IMAGE:4541304 5',
            mRNA sequence.
ACCESSION  BG334974
VERSION    BG334974.1  GI:13141412
KEYWORDS   EST.
SOURCE     human.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 704)
AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL   Unpublished (1999)
COMMENT    Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-re@mail.nih.gov
            Tissue procurement: ATCC
            cDNA library preparation: Ling Hong/Rubin Laboratory
            DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM1219 row: k column: 01
            High quality sequence stop: 699.

```

FEATURES

source

location/Qualifiers

1..704

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4541304"
/clone_lib="NIH_MGC_21"
/tissue_type="Choriocarcinoma"
/lab_host="DH10B (phage-resistant)"
/Note="Organ: placenta; Vector: pORF7; Site_1: XhoI;
       Site_2: EcoRI; cDNA made by oligo-dT priming.
       Directionally cloned into EcoRI/XhoI sites using the
       following 5' adaptor: GGCACGAG(G). Size-selected >500bp
       for average insert size 1.8kb. Library constructed by
       Ling Hong in the laboratory of Gerald M. Rubin (University
       of California, Berkeley) using ZAP-cDNA synthesis kit
       (Stratagene) and Superscript II RT (Life Technologies)."
```

BASE COUNT 114 a 240 c 229 g 121 t

alignment\_scores:
Quality: 1105.00 Length: 214
Ratio: 5.188 Gaps: 1
Percent Similarity: 99.533 Percent Identity: 99.533

alignment\_block:
US-09-525-867-1 x BG334974

Align seg 1/1 to: BG334974 from: 1 to: 704

```

1 MetaAlaValLeuSerAlaProGlyLeuArgGlyPheArgIleLeuGlyLe 17
|||||
4 ATGCGGTGCTGTGAGCTCTGGCTCGCGGCTTCCGGATCCTTGGTCT 53
|||||
17 uArgSerSerValGlyProAlaValGlnAlaArgGlyValHisGlnSerV 34
|||||
54 CGCTTCCAGCGTGGCGCGGCTGTGAGGACGAGGTGTCATCAGAGCG 103
|||||
34 alaIaThrAspGlyProSerSerThrGlnProAlaLeuProLysAlaArg 50
|||||
104 TGCCACCGATGCCCCAAGCAGCAGCCAGCTGCCCTGCCAAGGCCAGA 153
|||||
51 AlavaIaAlaProLysProSerArgGlyGluTyrValValAlaLysLe 67
|||||
154 GCGGTGGCTCCCAAGCCAGCAGCGGGCGAGTATGTGGTGGCCAAAGCT 203
|||||
67 uAspAspLeuValAsnTrpAlaArgArgSerSerLeuTrpProMetThrP 84
|||||
204 GGATGACCTGCTCACTGGCGCGCGGAGTCTCTGTGGCCCATGACCT 253
|||||
84 heGlyLeuAlaCysCysAlaValGluMetMetHisMetAlaAlaProArg 100
|||||
254 TCGGCTGGCTGCTGCGCGGTGGAGATGATGCATGGCAGCACCCCGC 303
|||||
101 TyrAspMetAspArgPheGlyValValPheArgAlaSerProArgGlnSe 117
|||||
304 TACGACATGACCGCTTGGCGTGGTCTTCCGCGCAGCGCCGCCAGTC 353
|||||
117 rAspValMetIleValAlaGlyThrLeuThrAsnLysMetAlaProAlaL 134
|||||
354 CGACGTCATGATCGTGGCGGCACACTCACCACAGATGGCCCGCAGCGC 403
|||||
134 euArgLysValTyrAspGlnMetProGluProArgTyrValValSerMet 150
|||||
404 TTCGAAGGCTACGACACAGATGCCGAGCGCGGTACGTGGTCTCCATG 453
|||||
151 GlySerCysAlaAsnGlyGlyTyrTyrHisTyrSerTyrSerValVa 167
|||||
454 GGGAGCTGCGCAACAGGAGGAGGCTACTACCATATTCCTACTCGTGTG 503
|||||
167 lArgGlyCysAspArgIleValProValAspIleTyrIleProGlyCys 183
|||||
504 GAGGGCTGGACCGCATCGTGGCGGTGGACATCTACATCCAGGCTGC 553
|||||
184 ProProThrAlaGluAlaLeuLeuTyrGlyIleLeuGlnLeuGlnArgLy 200

```

|||||  
 554 CCACCTACGGCCGAGCCCTGCTACGGCATCCCTGCAGCTGCAGAGAA 603  
 200 sileLysArgGluArgLeuGlnIleTrpTyrArgArg 213  
 |||||  
 604 GATCAAGCGGAGCGGAGGCTGCAGATCTGGTACCGCAGG 643

seq\_name: gb\_est2:BG706832

seq\_documentation\_block:

LOCUS BG706832 774 bp mRNA EST 07-MAY-2001  
 DEFINITION 602671984F1 NIH\_MGC\_96 Homo sapiens cDNA clone IMAGE:4794669 5',  
 mRNA sequence.

ACCESSION BG706832

VERSION BG706832.1 GI:13982569

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 774)

AUTHORS NIH-MGC <http://mgi.nci.nih.gov/>.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.

Email: [cgapbs-r@mail.nih.gov](mailto:cgapbs-r@mail.nih.gov)

Tissue Procurement: Miklos Palkovits, M.D., Ph.D.

cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki

Toshiyuki and Piero Carninci (RIKEN)

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

<http://image.llnl.gov>

Plate: LLAM10676 row: g column: 22

High quality sequence stop: 762.

FEATURES

source

1..774

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/clone="IMAGE:4794669"

/clone\_lib="NIH\_MGC\_96"

/tissue\_type="hypothalamus"

/lab\_host="DH10B"

/note="Organ: brain; Vector: pBluescriptR (modified

pBluescript KS+); Site\_1: BamHI; Site\_2: SalI-XhoI (gtcag

); Oligo-dT primed using primer 5'-TTTTTTTTTTTNN-3',

size-selected for average insert size 2.3 kb and

normalized to RCT 5. This is a primary library enriched

for full-length clones and constructed using the

Cap-trapper method (Carninci, in preparation). Library

constructed by M. Brownstein (NHGRI, National

Institutes of Health). Note: this is a NIH\_MGC library."

BASE COUNT

ORIGIN

alignment\_scores:

Quality: 1102.00 Length: 213

Ratio: 5.223 Gaps: 0

Percent Similarity: 99.061 Percent Identity: 99.061

alignment\_block:

US-09-525-867-1 x BG706832 ..

Align seg 1/1 to: BG706832 from: 1 to: 774

1 MetAlaValLeuSerAlaProGlyLeuArgGlyPheArgIleLeuGlyLe 17

|||||

33 ATGCGGGTGTGACCTCTGCTGGCTGGCGGCTTCCGGATCCCTTGCT 82

|||||

17 uArgSerValGlyProAlaValGlnAlaArgGlyValHisGlnSer 34

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

83 ACCTCCACCGCTGGGCGCGCTGTGCAGCAGAGGTGTCCATCAGACGC 132  
 34 alAlaThrAspGlyProSerSerThrGlnProAlaLeuProLysAlaArg 50  
 |||||  
 133 TGCCACCGATGTCGCCAACAGCAGCAGCCAGCTGCTGCCAAAGGCCAGA 182  
 51 AlaValAlaProLysProSerSerArgGlyGlyValValAlaLysLe 67  
 |||||  
 183 GCGGTGGCTCCCAACCCAGCAGCGCGGCGAGTATGGTGGCCAGCT 232  
 67 uAspAspLeuValAsnTrpAlaArgArgSerSerLeuTrpProMetThrP 84  
 |||||  
 233 GGATGACCTCGTCACTGGCGCCGCCGGAGTCTCTGTGGCCCATGACCT 282  
 84 heClyLeuAlaCysCysAlaValGluMetMetHisMetAlaAlaProArg 100  
 |||||  
 283 TCGGCTGGCTGCTGCGCGCTGGAGATGATGCATGCAGCAGCCACGC 332  
 101 TyrAspMetAspArgPheGlyValValPheArgAlaSerProArgGlnSe 117  
 |||||  
 333 TAGGACATGAGCGCTTTGGCGTGTCTTCCGCGCCAGCCGCCAGTC 382  
 117 rAspValMetIleValAlaGlyThrLeuThrAsnLysMetAlaProAlaL 134  
 |||||  
 383 CGACGTCATGATCGTGGCGCGCACACTCACCACCAAGATGGCGCAGCGC 432  
 134 euArgLysValTyrAspGlnMetProGluProArgTyrValValSerMet 150  
 |||||  
 433 TTGCGAAGTCTACGACCAAGATGGCGGCGCGCTAGCTGTCTCCATG 482  
 151 GlySerCysAlaAsnGlyGlyTyrTyrHisTyrSerTyrSerValVa 167  
 |||||  
 483 GGGAGTGGCGCCACGAGGAGGCTACTACCACTATTCTACTCGGTGT 532  
 167 largGlyCysAspArgIleValProValAspIleTyrIleProGlyCysP 184  
 |||||  
 533 GAGGGCTGCGACCGCATCGTCCCGTGGACATCTACATCCAGCTTGC 582  
 184 roProThrAlaGluAlaLeuLeuTyrGlyIleLeuGlnLeuGlnArgLys 200  
 |||||  
 583 CACCTACGCGCCAGCGCCCTGCTCTACGGCATCTCGACGTCGAGAGGAG 632  
 201 IleLysArgGluArgLeuGlnIleTrpTyrArgArg 213  
 |||||  
 633 ATCAACGGGAGCGGAGGCTGCAGATCTGGTACCGCAGG 671

seq\_name: gb\_est2:BG479937

seq\_documentation\_block:

LOCUS BG479937 836 bp mRNA EST 21-MAR-2001

DEFINITION 602527493F1 NIH\_MGC\_21 Homo sapiens cDNA clone IMAGE:4651103 5',

mRNA sequence.

ACCESSION BG479937

VERSION BG479937.1 GI:13412216

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 836)

AUTHORS NIH-MGC <http://mgi.nci.nih.gov/>.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.

Email: [cgapbs-r@mail.nih.gov](mailto:cgapbs-r@mail.nih.gov)

Tissue Procurement: ATCC

cDNA Library Preparation: Ling Hong/Rubin Laboratory

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

<http://image.llnl.gov>

Plate: LLCMI433 row: i column: 24



FEATURES High quality sequence stop: 751.  
Location/Qualifiers  
source

1. .836  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone\_lib="IMAGE:4651103"  
/clone\_lib="NIH\_MGC\_21"  
/tissue\_type="Choriocarcinoma"  
/lab\_host="DH10B (phage-resistant)"  
/note="Organ: placenta; Vector: pOTB7; Site\_1: XhoI;  
Site\_2: EcoRI; cDNA made by oligo-dT priming.  
Directionally cloned into EcoRI/XhoI sites using the  
following 5' adaptor: GGCACGAG(G). Size-selected >500bp  
for average insert size 1.8kb. Library constructed by  
Ling Hong in the laboratory of Gerald M. Rubin (University  
of California, Berkeley) using ZAP-cDNA synthesis kit  
(Stratagene) and Superscript II RT (Life Technologies)."  
BASE COUNT 152 a 281 c 261 g 141 t 1 others  
ORIGIN

alignment\_scores:  
Quality: 1102.00 Length: 214  
Ratio: 5.174 Gaps: 1  
Percent Similarity: 99.533 Percent Identity: 99.065  
alignment\_block:  
US-09-525-867-1 x BG479937 ..  
Align seg 1/1 to: BG479937 from: 1 to: 836

1 MetAlaValLeuSerAlaProGlyLeuArgGlyPheArgIleLeuGlyLe 17  
|||||  
7 ATGGCGGTGCTCAGCTCCTGGCTCGCGGCTTCGGATCCTTGGTCT 56  
|||||  
17 uArgSerSer.ValGlyProAlaValGlnAlaArgGlyValHisGlnSer 33  
|||||  
57 GCGTCCAGCGCTGGCGCGGCTGTGAGGACGAGGTGTCATCAGAGC 106  
|||||  
34 ValAlaThrAspGlyProSerSerThrGlnProAlaLeuProLysAla 50  
|||||  
107 GTGGCCACCGATGGCCCAAGCAGCACCGCCCTGCTGCCAAAGGCCAG 156  
|||||  
50 GaLaValAlaProLysProSerSerArgGlyGluTyrValValAlaLysL 67  
|||||  
157 AGCGGTGGCTCCCAACCCAGCAGCGGGCGAGATGTGTGGCCCAAGC 206  
|||||  
67 euAspAspLeuValAsnTrpAlaArgArgSerSerLeuTrpProMetThr 83  
|||||  
207 TGGATGACCTGCTCACTGGCCCGCGGAGTTCTCTGTGGCCCATGACC 256  
|||||  
84 PheGlyLeuAlaCysCysAlaValGluMetMetHisMetAlaAlaProAr 100  
|||||  
257 TTCGGCTGGCTGCTGGCGCGTGAGATGATGCATGCGCAGCACCCCG 306  
|||||  
100 gTyrAspMetAspArgPheGlyValValPheArgAlaSerProArgGln 117  
|||||  
307 CTACGACATGACCGCTTTGGCGTGGTCTTCCGGCCAGCCCGCGCAGT 356  
|||||  
117 erAspValMetIleValAlaGlyThrLeuThrAsnLysMetAlaProAla 133  
|||||  
357 CCGAGTGTATGATGCTGGCGCGCACACTCACCACCAAGATGGCCCGAG 406  
|||||  
134 LeuArgLysValTyrAspGlnMetProGluProArgTyrValValSerMe 150  
|||||  
407 CTTGCCAAGGCTACGACCATGCGGAGCCCGCGCTACGTGCTCCAT 456  
|||||  
150 tClySerCysAlaAsnGlyGlyTyrHisTyrSerTyrSerValV 167  
|||||  
457 GGGAGGCTGCGCCCAACGAGGAGGCTACTACCACTATTCCTACTCGGTG 506  
|||||  
167 aLArgGlyCysAspArgIleValProValAspIleTyrIleProGlyCys 183  
|||||

507 TGAGGGCTGCGACCGCATCTGCCGTGGACATCTACATCCAGGCTGC 556  
184 ProProThrAlaGluAlaLeuLeuTyrGlyLeuGlnLeuGlnArgTy 200  
|||||  
557 CCACCTACGCGCGAGGCCCTGCTCTACGGCATCTCGAGCTCGAGAGAA 606  
200 sIleLysArgGluArgArgLeuGlnIleTrpTyrArgArg 213  
|||||  
607 GATCAACGGGAGCGGAGGCTGCAGATCTGGTACCGCAGG 646  
seq\_name: gb\_est2:BF342354  
seq documentation\_block:  
LOCUS BF342354 804 bp mRNA EST 22-NOV-2000  
DEFINITION 602013145F1 NCI\_CGAP\_Brn64 Homo sapiens cDNA clone IMAGE:4148842  
5', mRNA sequence.  
ACCESSION BF342354  
VERSION BF342354.1 GI:11289355  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 804)  
AUTHORS NIH-MGC http://mgi.nci.nih.gov/  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgabbs@mail.nih.gov  
Tissue Procurement: David N. Louis, M.D.  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: LLAM9409 row: j column: 11  
High quality sequence start: 2  
High quality sequence stop: 774.  
FEATURES  
source  
1. .804  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone\_lib="NCI\_CGAP\_Brn64"  
/tissue\_type="glioblastoma with EGFR amplification"  
/lab\_host="DH10B (T1 phage-resistant)"  
/note="Organ: brain; Vector: pCMV-SPORT6; Site\_1: NotI;  
Site\_2: SalI; Cloned unidirectionally. Primer: Oligo dT.  
Average insert size 1.57 kb. Constructed by Life  
Technologies. Note: this is a NCI\_CGAP Library."  
BASE COUNT 149 a 268 c 255 g 132 t  
ORIGIN

alignment\_scores:  
Quality: 1081.00 Length: 211  
Ratio: 5.123 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000  
alignment\_block:  
US-09-525-867-1 x BF342354 ..  
Align seg 1/1 to: BF342354 from: 1 to: 804  
1 MetAlaValLeuSerAlaProGlyLeuArgGlyPheArgIleLeuGlyLe 17  
|||||  
23 ATGGCGGTGCTCAGCTCCTGGCTCGCGGCTTCGGATCCTTGGTCT 72  
|||||  
17 uArgSerSerValGlyProAlaValGlnAlaArgGlyValHisGlnSerV 34  
|||||  
73 GCGTCCAGCGTGGCGCGGCTGTGAGGACGAGGTGTCATCAGAGCG 122  
|||||



```

|||||
559 ACGGCATCTCTCAGCTGCAGAGGAGATCAAGCGGAGCGGAGCGTGCAG 608
209 llettrpTyArgArg 213
|||||
609 ATCTGGTACCGCAGG 623

```

seq\_name: gb\_est2:BE794266

```

seq_documentation_block:
LOCUS BE794266 879 bp mRNA EST 20-SEP-2000
DEFINITION 60159151F1 NIH_MGC_7 Homo sapiens cDNA clone IMAGE:3945589 5',
mRNA sequence.
ACCESSION BE794266
VERSION BE794266
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 879)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs@mail.nih.gov
Tissue Procurement: DCTD/DTP
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
Plate: LLC804 row: e column: 14
High quality sequence stop: 773.
Location/Qualifiers
1..879
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3945589"
/clone_lib="NIH_MGC_7"
/tissue_type="small cell carcinoma"
/cell_line="MGC3"
/lab_host="DH10B (phage-resistant)"
/note="Organ: lung; Vector: pOTB7; Site: 1: XhoI; Site 2:
EcoRI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGCACGAG(S). Size-selected >500bp for average
insert size 1.8kb. Library constructed by Ling Hong in
the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."
```

```

BASE COUNT 166 a 276 c 280 g 157 t
ORIGIN
alignment_scores:
Quality: 1080.00 Length: 213
Ratio: 5.118 Gaps: 0
Percent Similarity: 99.061 Percent Identity: 99.061
alignment_block:
US-09-525-867-1 x BE794266 ..
Align seg 1/1 to: BE794266 from: 1 to: 879
1 MetAlaValLeuSerAlaProGlyLeuArgGlyPheArgIleuLeuGlyLe 17
22 ATGGCGGTGCTGTGAGTCTCTGCGCGGCTCCGATCTCGGTCT 71
17 uArgSerSerValGlyProAlaValGlnAlaArgGlyValHisGlnSerV 34
72 GCGTCCAGCGTGGCGCTGCTGTGCAGGCACGAGGTGTCCATCAGAGCG 121

```

```

34 alAlaThrAspGlyProSerSerThrGlnProAlaLeuProLysAlaArg 50
|||||
122 TGGCCACCGATGGCCCAAGCAGCAGCCAGCTGCCCTGCCAAAGCCAGA 171
51 AlaValAlaProLysProSerSerArgGlyGlyTyrValValAlaLysLe 67
|||||
172 GCGGTGGCTCCCAAAACCAGCAGCCGGGGGAGTATGTGTGGTGGCCAGCT 221
57 uAspAspLeuValAsnTrpAlaArgArgSerSerLeuTrpProMetThrP 84
|||||
222 GGATGACCTGCTCAACTGGGCGCGCGGAGTTCTCTGTGGCCCATGACCT 271
84 heGlyLeuAlaCysCysAlaValGluMetMetHisMetAlaAlaProArg 100
|||||
272 TCGGCTGGCTGCTGCGCGTGGAGTATGCATGGCAGCACCCCGC 321
101 TyrAspMetAspArgPheGlyValValPheArgAlaSerProArgGlnSe 117
|||||
322 TACGACATGGACCGCTTTGGCTGGTCTCCGCGCCAGCCGCGCCAGTC 371
117 rAspValMetIleValAlaGlyThrLeuThrAsnLysMetAlaProAlaL 134
|||||
372 CGAGCTCATGATCGTGGCGCGCACACTCACCACACAGATGGCCCCAGCGC 421
134 euArgLysValTyrAspGlnMetProGluProArgTyrValValSerMet 150
|||||
422 TTCGCAAGGCTACGACACAGATCGCGAGCGCGCTACGTGGTCTCCATG 471
151 GlySerCysAlaAsnGlyGlyTyrTyrHisTyrSerTyrSerValVa 167
|||||
472 GGGAGCTGGCCCAACGAGGAGGCTACTACCACTATCTCTACTCGGTGT 521
167 lArgGlyCysAspArgIleValProValAspIleTyrIleProGlyCysP 184
|||||
522 GAGGGGTGGCAGCG.ATCGTGGCGTGGACATCTACATCCAGGCTGCC 570
184 roPcThrAlaGluAlaLeuTyrGlyIleLeuGlnLeuGlnArgLys 200
|||||
571 CACCTAGCGCGGAGGCGCTCTAGGCGCATCTCGAGCTGCAGAGGAAG 620
201 lIeLysArgGluArgArgLeuGlnIleTrpTyrArgArg 213
621 ATCAAGCGGAGGAGGCGCTGCAGATCTGTGTACCGCAGG 659

```

seq\_name: gb\_est2:BG491103

```

seq_documentation_block:
LOCUS BG491103 738 bp mRNA EST 27-MAR-2001
DEFINITION 602519337F1 NIH_MGC_18 Homo sapiens cDNA clone IMAGE:4638441 5',
mRNA sequence.
ACCESSION BG491103
VERSION BG491103.1 GI:13452615
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 738)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs@mail.nih.gov
Tissue Procurement: DCTD/DTP/Gazdar
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM1400 row: j column: 10
High quality sequence stop: 738.
Location/Qualifiers

```

FEATURES

```

source
1. 738
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4638441"
/clone_lib="NIH_MGC_18"
/tissue_type="large cell carcinoma"
/lab_host="DH10B (phage-resistant)"
/notes="Organ: lung; Vector: pOTB7; Site_1: XhoI; Site_2:
EcoRI; cDNA made by oligo-dt priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGGAG(G). Library constructed by Ling Hong in the
Laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
Superscript II RT (Life Technologies). Note: this is a
NIH_MGC Library."
BASE COUNT 125 a 252 c 235 g 126 t
ORIGIN

alignment_scores:
Quality: 1068.00 Length: 211
Ratio: 5.110 Gaps: 0
Percent Similarity: 99.052 Percent Identity: 99.052

alignment_block:
US-09-525-867-1 x BG491103
Align seg 1/1 to: BG491103 from: 1 to: 738

1 MetAlaValLeuSerAlaProGlyLeuArgGlyPheArgIleLeuGlyLe 17
|||||
7 ATGGCGGTGCTGCTCAGCTCTGCGCTGCGCGCTTCCGGATCCTTGCT 56
|||||
17 uArgSerValGlyProAlaValGlnAlaArgGlyValHisGlnServ 34
|||||
57 GCCTCCAGCTGGGCTGGCTGTGACGACGAGGTGTCCATCAGACG 106
|||||
34 alaThrAspGlyProSerSerThrGlnProAlaLeuProLysAlaArg 50
|||||
107 TGCCACCGATGCCCAAGCAGCAGCCAGCCTGCCCTGCCAAGSCCAGA 156
|||||
51 AlaValAlaProLysProSerSerArgGlyGluTyValValAlaLysLe 67
|||||
157 GCGGTGGCTCCCAACCCAGCAGCGGGGCGAGTATGGTGGCCAAAGCT 206
|||||
67 uAspAspLeuValAsnTrpAlaArgArgSerSerLeuTrpProMetThr 84
|||||
207 GGATGACCTCGTCACTGGCCCGCGGAGTCTCTGTGGCCATGACCT 256
|||||
84 heGlyLeuAlaCysCysAlaValGluMetMetHisMetalAlaProArg 100
|||||
257 TCGGCTGGCTGCTGCTGCGCGGTGGAGATGATGCATGGCAGCAGCCGC 306
|||||
101 TyrAspMetAspArgPheGlyValValPheArgAlaSerProArgGlnSe 117
|||||
307 TAGGACATGACCGCTTTGGCGTGGTCTTCGCGCCAGCCGCGCAGTC 356
|||||
117 rAspValMetIleValAlaGlyThrLeuThrAsnLysMetalAlaProAla 134
|||||
357 CGACGTCATGATCGTGGCGGCACACTCACCACCAAGATGGCCCGCAGCG 406
|||||
134 euArgLysValTyAspGlnMetProGluProArgTyValValSerMet 150
|||||
407 TTCGAAGGCTTACGACCAAGATGCCGAGCCGCGTACGTGCTCCATG 456
|||||
151 GlySerCysAlaAsnGlyGlyTyTrpHisTySerTySerValVal 167
|||||
457 GGGAGCTGGCCCAAGCAGGAGGCTACTACCACTATTCCTACTCGTGTGT 506
|||||
167 lArgGlyCysAspArgIleValProValAspIleTyrlleProGlyCysP 184
|||||
507 GAGGGGCTGGACCGATGTCGCCGTGGACATCTACATCCAGGCTGCC 555

```

```

184 roProThrAlaGluAlaLeuLeuTyGlyIleLeuGlnLeuGlnArgLys 200
|||||
556 A.CCTACGCCGAGGCGCTCTACGGCATCTCGAGCTGCAGAGGAAG 605
|||||
201 IleLysArgGluArgArgGluGlnIleTrpTyr 211
|||||
606 ATCAACGGGAGGAGGAGGCTGCAGATCTGGTAC 638
|||||
seq_name: gb_est2:B116730
seq_documentation_block:
LOCUS B116730 733 bp mRNA EST 26-JUN-2001
DEFINITION 602868418F1 NIH_MGC_7 Homo sapiens cDNA clone IMAGE:5017179 5',
mRNA sequence.
ACCESSION B116730
VERSION B116730.1 GI:14567631
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 733)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs@mail.nih.gov
Tissue Procurement: DCTD/DTP
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
Plate: L1CM1828 row: 0 column: 04
High quality sequence stop: 670.
Location/Qualifiers
1..733
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5017179"
/clone_lib="NIH_MGC_7"
/tissue_type="small cell carcinoma"
/lab_host="DH10B (phage-resistant)"
/notes="Organ: lung; Vector: pOTB7; Site_1: XhoI; Site_2:
EcoRI; cDNA made by oligo-dt priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGCAGGAG(G). Size-selected >500bp for average
insert size 1.8kb. Library constructed by Ling Hong in
the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).".
BASE COUNT 130 a 253 c 230 g 120 t
ORIGIN

alignment_scores:
Quality: 1062.00 Length: 204
Ratio: 5.232 Gaps: 0
Percent Similarity: 99.510 Percent Identity: 99.020

alignment_block:
US-09-525-867-1 x B116730
Align seg 1/1 to: B116730 from: 1 to: 733

10 ArgGlyPheArgIleLeuGlyLeuArgSerValGlyProAlaValGln 26
|||||
3 CGCGGCTCCGGATCCTTGGCTGCGCTCCAGCTGGCGCTGGCTGTGCA 52
|||||
26 nAlaArgGlyValHisGlnSerValAlaThrAspGlyProSerSerThr 43
|||||
53 GGCACGAGGTGCTCCATCAGAGCGTGGCCACCGATGGCCCAAGCAGCACC 102
|||||

```

```

43  lnProAlaLeuProLysAlaArgAlaValAlaProLysProSerSerArg 59
|||||
103 AGCCTGCCCTGCCAAGCCAGAGCGCTGGCTCCCAACCCAGAGCGCG 152
|||||
60  GlyGluTyrValValAlaLysLeuAspLeuValAlaSerTrpAlaArgAr 76
|||||
153 GCGAGTATGTGGTGGCCAAAGCTGGATGACCTGCTCAACTGGGCGCCG 202
|||||
76  gSerSerLeuTrpProMetThrPheGlyLeuAlaCysCysAlaValGluM 93
|||||
203 GAGTCTCTGTGGCCCATGACTTGGGCTGGCTGGCTGGCTGGCTGGAGA 252
|||||
93  etMetHisMetAlaAlaProArgTyrAspMetAspArgPheGlyValVal 109
|||||
253 TGATGCACATGCGACGACCCCGCTACGACATGGACCGCTTTGGGCTGTC 302
|||||
110 PheArgAlaSerProArgGlnSerAspValMetIleValAlaGlyThrLe 126
|||||
303 TTCCGCGCCAGCCCGCCGACGTCGACGTCATGATCGTGGCCGSCACACT 352
|||||
126 uThrAsnLysMetAlaProAlaLeuArgLysValTyrAspGlnMetProG 143
|||||
353 CACCAACAGATGGCCCGCAGCGCTTCGCAAGTCTACGACAGATGCCGG 402
|||||
143 luProArgTyrValValSerMetGlySerCysAlaAsnGlyGlyTyr 159
|||||
403 AGCCGCGCTACGTGCTCCATGGGAGCTGCGCCAAACGGAGGAGGTAC 452
|||||
160 TyrHisTyrSerTyrSerValValArgGlyCysAspArgIleValProVa 176
|||||
453 TACCACTATCTCTACTCGTGGTGGGGGCTGGACCGCATCGTGGCCGT 502
|||||
176 lAspIleTyrIleProGlyCysProProThrAlaGluAlaLeuLeuTyrG 193
|||||
503 GGACATCATACATCCAGCGCTGCCACCTACGTGGAGGCGCTGCTCTACG 552
|||||
193 lYlleLeuGlnLeuGlnArgLysIleLysArgGluArgArgLeuGlnIle 209
|||||
553 GCATCTCAGCTGCAGAGGAAGATCAAGCGGGAGCGGAGGCTGCAGATC 602
|||||
210 TrpTyrArgArg 213
|||||
603 TGGTACCGCAGG 614

```

seq\_name: gb\_est2:BI195871

```

seq_documentation_block:
LOCUS      BI195871          788 bp    mRNA          EST          10-JUL-2001
DEFINITION 602756115F1 NIH_MGC_19 Homo sapiens cDNA clone IMAGE:4891946 5',
            mRNA sequence.
ACCESSION  BI195871
VERSION    BI195871.1 GI:14650891
KEYWORDS   EST.
SOURCE     human.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 788)
            NIH-MGC http://mgs.nci.nih.gov/.
            National Institutes of Health, Mammalian Gene Collection (MGC)
            Unpublished (1999)
            Contact: Robert Strausberg, Ph.D.
            Email: cgabbs-r@mail.nih.gov
            Tissue Procurement: ATCC

```

```

CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
Plate: LLCMI771 row: m column: 03
High quality sequence stop: 770.
Location/Qualifiers

```

FEATURES

```

source      1..788
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /clone_lib="IMAGE:4891946"
            /clone_lib="NIH_MGC_19"
            /tissue_type="neuroblastoma"
            /lab_host="DH10B (phage-resistant)"
            /note="Organ: brain; Vector: pOTB7; Site_1: XhoI; Site_2:
            EcoRI; cDNA made by oligo-dT priming. Directionally
            cloned into EcoRI/XhoI sites using the following 5'
            adaptor: GGCACGAG(G). Library constructed by Ling Hong
            in the laboratory of Gerald M. Rubin (University of
            California, Berkeley) using ZAP-cDNA synthesis kit
            (Stratagene) and Superscript II RT (Life Technologies).
            Note: this is a NIH_MGC library."
BASE COUNT  147 a 254 c 254 g 133 t
ORIGIN

```

```

alignment_scores:
Quality: 1059.50      Length: 212
Ratio: 5.094          Gaps: 2
Percent Similarity: 98.113 Percent Identity: 97.642
alignment_block:
US-09-525-867-1 x BI195871
Align seg 1/1 to: BI195871 from: 1 to: 788

```

```

3  ValLeuSerAlaProGlyLeuArgGlyPheArgIleLeuGlyLeuArgSe 19
|||||
29  GTGCTGTGTCAGCTCTGGCTGCGGGCTTCGGATCCTTGTGTGGCTC 78
|||||
19  rSerValGlyProAlaValGlnAlaArgGlyValHisGlnSerValAla 36
|||||
79  CAGCGTGGCCCGCGCTGTGACGACACGAGGTGTCCATCAGAGCGTGGCCA 128
|||||
36  hrAspGlyProSerSerThrGlnProAlaLeuProLysAlaArgAlaVal 52
|||||
129 CCGATGGCCCAAGCAGCAGCAGCGCTGCCCTGCCAAAGCCAGAGCGGTG 178
|||||
53  AlaProLysProSerSerArgGlyGluTyrValValAlaLysLeuAsp 69
|||||
179 GCTCCCAAAACCCAGCAGCGGGCGAGTATGTGTGGCCCAAGCTGGATGA 228
|||||
69  pLeuValAsnTrpAlaArgArgSerSerLeuTrpProMetThrPheGlyL 86
|||||
229 CCGTCAACTGGCCCGCGCGAGTTCCTGTGGCCCATGACCTTCGGCC 278
|||||
86  euAlaCysCysAlaValAlaGluMetMetHisMetAlaAlaProArgTyr 102
|||||
279 TGGCGCTGCTGGCGCTGGAGATGATGCACATGGCAGCACCCCGCTACGAC 328
|||||
103 MetAspArgPheGlyValValPheArgAlaSerProArgGlnSerAspVa 119
|||||
329 ATGGACCGCTTTGGCTGGTCTTCCGCGCCAGCCGCGCCAGTCCGACGT 378
|||||
119 lMetIleValAlaGlyThrLeuThrAsnLysMetAlaProAlaLeuArgL 136
|||||
379 CATGATCGTGGCCCGCACACTCACCAAGATGGCCCGCCAGCGCTTCGCA 428
|||||
136 ysValTyrAspGlnMetProGluProArgTyrValValSerMetGlySer 152
|||||
429 AGGTCTACGACCATGCGCGAGCGCGCTACGTGGTCTCCATGGGGAGC 478
|||||
153 CysAlaAsnGlyGlyTyrTyrHisTyrSerTyrSerValValArgGl 169
|||||
479 TGCGCCAACGGAGAGGCTACTACCACTATTCCTACTCGGTGGTGGGG 528
|||||
169 yCysAspArgIleValProValAspIleTyrIleProGlyCys.ProPro 185
|||||
529 CTGCGACCGCATCGTGGCGGTGGACATCTACATCCCGAGGCTGCCACC... 575

```



```

/tissue_type="small cell carcinoma"
/cell_line="MGC3"
/lab_host="DH10B (phage-resistant)"
/notes="Organ: lung; Vector: pOTB7; Site_1: XhoI; Site_2:
EcoRI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGCACGAG(G). Size-selected >500bp for average
insert size 1.8kb. Library constructed by Ling Hong in
the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."
```

BASE COUNT 127 a 244 c 240 g 125 t  
 ORIGIN

alignment\_scores:  
 Quality: 1054.00 Length: 213  
 Ratio: 4.995 Gaps: 0  
 Percent Similarity: 99.061 Percent Identity: 99.061

alignment\_block:  
 US-09-525-867-1 x BE795373

Align seg 1/1 to: BE795373 from: 1 to: 736

```

1 MetAlaValLeuSerAlaProGlyLeuArgGlyPheArgIleLeuGlyLe 17
|||||
8 ATGGGGGTGCTGTCAGCTCCCTGGCTG.CGCGGCTTCGGGATCCTGGTCT 56
|||||
17 uArgSerSerValGlyProAlaValGlnAlaArgGlyValHisGlnSerV 34
|||||
57 GCGCTCCAGCTGGGCTGCTGTCAGGACGAGGTGCTCATCAGAGCG 106
|||||
34 alAlaThrAspGlyProSerSerThrGlnProAlaLeuProLysAlaArg 50
|||||
107 TGGCCACCGATGGCCCAAGCAGCACCCAGCTGCCCTGCCCAAGGCCAGA 156
|||||
51 AlaValAlaProLysProSerSerArgGlyGluTyrValValAlaLysLe 67
|||||
157 GCGGTGGCTCCCAACCCAGCAGCGGGGGAGTATGTGTGGCCCAAGCT 206
|||||
67 uAspAspLeuValAsnTrpAlaArgArgSerSerLeuTrpProMetThrP 84
|||||
207 GGATGACCTGCTCAACTGGGCCCCCGGAGTTCTCTGTGGCCCATGACCT 256
|||||
84 heGlyLeuAlaCysCysAlaValGluMetMethHisMetAlaAlaProArg 100
|||||
257 TGGGCTGCTGCTGCGCCGTGGAGATGATGCACATGGCAGCACCCCGC 306
|||||
101 TyrAspMetAspArgPheGlyValValPheArgAlaSerProArgGlnSe 117
|||||
307 TAGGACATGGACCGCTTGGCGTGGTCTTCGCGCCAGCCCGCCGCGATC 356
|||||
117 rAspValMetIleValAlaGlyThrLeuThrAsnLysMetAlaProAlaL 134
|||||
357 CGAGCTCATGATGCTGCGCCGACACATCACCAAGATGGCCCGCCAGCG 406
|||||
134 euArgLysValTyrAspGlnMetProGluProArgTyrValValSerMet 150
|||||
407 TTCGCAAGTCTAGACCAAGTGGCGGAGCGCGCTACGTGGTCTCATG 456
|||||
151 GlySerCysAlaAsnGlyGlyTyrTyrHisTyrSerTyrSerValVal 167
|||||
457 GGGAGCTCGGCCAACGGAGGAGGCTACTACCACTATTCCTACTCGGTGGT 506
|||||
167 lArgGlyCysAspArgIleValProValAspIleTyrIleProGlyCysP 184
|||||
507 GAGGGCTGCGACCGATCGTGGCCGGGACATCTACATCCCGAGGTGCC 556
|||||
184 roProThrAlaGluAlaLeuLeuTyrGlyIleLeuGlnLeuAlaArgLys 200
|||||
557 CACCTACGGC.GAGGCCCTGCTCTACGGCATCTCGAGCTGCAGAGGGAAG 605
|||||
```

```

201 IleLysArgGluArgArgLeuGlnIleTrpTyrArgArg 213
|||||
606 ATCAGCGGGAGCGAGGCTGTCAGATCTGTGTCAGCGAGG 644
seq_name: gb_est2:BG419411
seq_documentation_block:
  LOCUS BG419411 910 bp mRNA EST 14-MAR-2001
  DEFINITION 602446102F1 NTH_MGC_14 Homo sapiens cDNA clone IMAGE:4584528 5',
  mRNA sequence.
  ACCESSION BG419411
  VERSION BG419411.1 GI:13325917
  KEYWORDS EST
  SOURCE human.
  ORGANISM Homo sapiens
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
  REFERENCE 1 (bases 1 to 910)
  AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
  TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL Unpublished (1999)
  COMMENT Contact: Robert Strausberg, Ph.D.
  Email: cgabbs-remail.nih.gov
  Tissue Procurement: DCTD/DTF
  CDNA Library Preparation: Ling Hong/Rubin Laboratory
  CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
  DNA Sequencing by: Incyte Genomics, Inc.
  Clone Distribution: MGC clone distribution information can be
  found through the I.M.A.G.E. Consortium/LLNL at:
  http://image.llnl.gov
  Plate: LLCMI310 row: d column: 01
  High quality sequence stop: 702.
  FEATURES
    Location/Qualifiers
      1..910
        organism="Homo sapiens"
        gb_xref="taxon:9606"
        clone="IMAGE:4584528"
        tissue_type="renal cell adenocarcinoma"
        lab_host="DH10B (phage-resistant)"
        note="Organ: Kidney; Vector: pOTB7; Site_1: XhoI; Site_2:
        EcoRI; cDNA made by oligo-dT priming. Directionally
        cloned into EcoRI/XhoI sites using the following 5'
        adaptor: GGCACGAG(G). Size-selected >500bp for average
        insert size 1.8kb. Library constructed by Ling Hong in
        the laboratory of Gerald M. Rubin (University of
        California, Berkeley) using ZAP-cDNA synthesis kit
        (Stratagene) and Superscript II RT (Life Technologies)."
```

```

BASE COUNT 195 a 279 c 278 g 156 t
ORIGIN

alignment_scores:
  Quality: 1051.50 Length: 211
  Ratio: 5.104 Gaps: 1
  Percent Similarity: 97.630 Percent Identity: 96.682

alignment_block:
  US-09-525-867-1 x BG419411
  Align seg 1/1 to: BG419411 from: 1 to: 910
  1 MetAlaValLeuSerAlaProGlyLeuArgGlyPheArgIleLeuGlyLe 17
  |||||
  8 ATGGCGGTGCTGTCAGCTCCCTGGCTGCGCGGCTTCGGGATCCTGGTCT 57
  |||||
  17 uArgSerSerValGlyProAlaValGlnAlaArgGlyValHisGlnSerV 34
  |||||
  58 GCGCTCCAGCGTGGCGCTGCTGTGCAGGCAGCGAGGTGCTCATCAGAGCG 107
  |||||
  34 alAlaThrAspGlyProSerSerThrGlnProAlaLeuProLysAlaArg 50
  |||||
  108 TGGCCACCGATGGCCCAAGCAGCACCCAGCTGCTGCCCAAGGCCAGA 157
  |||||
```

```

51 AlaValAlaLeuProLysProSerSerArgGlyGlyTyrValValAlaLysLe 67
|||||
158 GCCGTGCTCCCAAAACCCAGACCGCGCGAGTATGTGTGGCCCAAGCT 207
|||||
67 uAspAspLeuValAsnTrpAlaArgSerSerLeuTrpProMetThrP 84
|||||
208 GGATGACCTCGTCAACTGGCCCGCGGAGTCTCTGTGGCCCATGACCT 257
|||||
84 heGlyLeuAlaCysCysAlaValGluMetMetHisMetAlaAlaProArg 100
|||||
258 TCGGCCTGGGCTGCTGGCCGCGTGAGATGACATGGCAGCACCCCGC 307
|||||
101 TyrAspMetAspArgPheGlyValValPheArgAlaSerProArgGlnSe 117
|||||
308 TACGACATGACCGCTTTGGCGTGGTCTTCCGGCCGAGCCCGCGCAGTC 357
|||||
117 rAspValMetIleValAlaGlyThrLeuThrAsnLysMetAlaProAlaL 134
|||||
358 CGAGCTCATGATCGTGGCCGGCACACTCACCAACAAGATGGCCCGAGC 407
|||||
134 euArgLysValTyrAspGlnMetProGluProArgTyrValValSerMet 150
|||||
408 TTCGCAAGGTCTACGACACAGATCCGAGCGCGCTACGTGTCTCCATG 457
|||||
151 GlySerCysAlaAsnGlyGlyTyrTyrHisTyrSerTyrSerValVa 167
|||||
458 GGGAGCTGGCCCAACGAGGAGGCTACTACCACTATTCCTACTCGTGTG 507
|||||
167 lArgGlyCysAspArgIleValProValAspIleTyrIleProGlyCysP 184
|||||
508 GAGGGGTGGCGCGCATCGTGGCCGTGGACATCTACATCCCGAGCTGC. 556
|||||
184 roProThrAlaGluAlaLeuLeuTyrGlyIleLeuGlnLeuGlnArgLys 200
|||||
557 ..CACCTACCGCGAGGCTGCTCTACGGCATCTCGAGCTGAGAGGAAG 604
|||||
201 lIleLysArgGluArgArgLeuGlnIleTrpTyr 211
|||||
605 ATCAAGCGGAGGAGGCTGCAGATCTGTGATC 637
|||||

seq_name: gb_est2:BG395342

seq_documentation_block:
LOCUS BG395342 1002 bp mRNA EST 12-MAR-2001
DEFINITION 6024578791 NIH_MGC_16 Homo sapiens cDNA clone IMAGE:4580538 5',
mRNA sequence.
ACCESSION BG395342
VERSION BG395342.1 GI:13288790
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue procurement: ATCC
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
Plate: LLCM1299 row: m column: 19
High quality sequence stop: 753.
Location/Qualifiers
1..1002
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4580538"

```

```

/clone_lib="NIH_MGC_16"
/tissue_type="retinoblastoma"
/lab_host="DH10B (phage-resistant)"
/Note="Organ: eye; Vector: pOTB7; Site:1: XhoI; Site_2:
EcoRI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGCACGAG(G). Library constructed by Ling Hong
in the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."

```

```

BASE COUNT      215 a      313 c      321 g      153 t
ORIGIN

```

```

alignment_scores:
  Quality: 1051.00      Length: 216
  Ratio: 5.005          Gaps: 3
  Percent Similarity: 97.222      Percent Identity: 97.222

```

```

alignment_block:
US-09-525-867-1 x BG395342

```

```

Align seg 1/1 to: BG395342 from: 1 to: 1002

```

```

1 MetAlaValLeuSerAlaProGlyLeuArgGlyPheArgIleLeuGlyLe 17
|||||
20 ATGGCGGTGCTGTACAGTCTCTGGCCTCGCGGCTTCCGGATCCTTGGTCT 69
|||||
17 uArgSerSerValGlyProAlaValGlnAlaArgGlyValHisGlnSerV 34
|||||
70 GCGCTCAGCGTGGGCTGGCTGTGCAGGCAGCAGGTGTCCATCAGAGCG 119
|||||
34 alAlaThrAspGlyProSerSerThrGlnProAlaLeuProLysAlaArg 50
|||||
120 TGGCACCGATGGCCCAAGCAGCAGCCCGCTGCCCTGCCAAAGGCCAGA 169
|||||
51 AlaValAlaProLysProSerSerArgGlyGlyTyrValValAlaLysLe 67
|||||
170 GCGTGTGCTCCCAACCCAGCAGCGCGGCGAGTATGTGTGGCCCAAGCT 219
|||||
67 uAspAspLeuValAsnTrpAlaArgSerSerLeuTrpProMetThrP 84
|||||
220 GGATGATCTCGTCAACTGGCGCGCGCGAGTCTCTGTGGCCCATGACCT 269
|||||
84 heGlyLeuAlaCysCysAlaValGluMetMetHisMetAlaAlaProArg 100
|||||
270 TCGGCTGGGCTGCTGCGCGGTGGAGATGATGCATGTCACACCCCGCC 319
|||||
101 TyrAspMetAspArgPheGlyValValPheArgAlaSerProArgGlnSe 117
|||||
320 TAGCAGATGGACCGCTTGGCGTGTCTTCCGCGCCAGCCCGCGCAGTC 369
|||||
117 rAspValMetIleValAlaGlyThrLeuThrAsnLysMetAlaProAlaL 134
|||||
370 CGAGCTCATGATCGTGGCGGCACACTCACCAACAAGATGGCCCGCCAG 419
|||||
134 euArgLysValTyrAspGlnMetProGluProArgTyrValValSerMet 150
|||||
420 TTCGCAAGGTCTACGACACAGATCCCGGAGCGCGCTACGTGTCTCCATG 469
|||||
151 GlySerCysAlaAsnGlyGlyTyrTyrHisTyrSerTyrSerValVa 167
|||||
470 GGGAGCTGGCCCAACGAGGAGGCTACTACCACTATTCCTACTCGTGTGT 519
|||||
167 lArgGlyCysAspArgIleValProValAspIleTyrIleProGlyCysP 184
|||||
520 GAGGGCTGGACCGCATCGTGGCCGTGGACATCTACATCCCAAGGCTGC 569
|||||
184 roProThrAlaGluAlaLeuLeuTyrGlyIleLeuGlnLeuGlnArg 199
|||||
570 CACCTACGCGCGAGGCGCTGTCTACGGCATTCCTGCAGGTTGCAGAGG 619
|||||

```



200 LysileLysArgGluArgGluGlnIleTyrArgArg 213  
|||||  
620 AAGATCAAGCGGAGGAGGAGGCTGAGATCTGTACCGCAGG 661

seq\_name: gb\_est2.BG765790

seq\_documentation\_block: 742 bp mRNA EST 15-MAY-2001  
LOCUS BG765790 NIH\_MGC\_49 Homo sapiens cDNA clone IMAGE:4869532 5',  
DEFINITION 602738620F1  
mRNA sequence.

ACCESSION BG765790  
VERSION BG765790.1 GI:14076443  
KEYWORDS EST.  
SOURCE human.

ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 742)  
NIH-MGC <http://mgc.nci.nih.gov/>.  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-r@mail.nih.gov

Tissue Procurement: ATCC/DCTD/DTP  
cDNA Library Preparation: Ling Hong/Rubin Laboratory  
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>

Plate: LICM1741 row: g column: 05  
High quality sequence stop: 739.  
Location/Qualifiers

FEATURES  
source

1. 742  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:4869532"  
/clone\_lib="NIH\_MGC\_49"  
/tissue\_type="melanotic melanoma, high MDR (cell line)"  
/lab\_host="DH10B (phage-resistant)"  
/note="Organ: skin; Vector: pOTB7; Site\_1: XhoI; Site\_2:  
EcoRI; cDNA made by oligo-dt priming. Directionally cloned  
into EcoRI/XhoI sites using the following 5' adaptor:  
GGCAGCAG(G). Size-selected >500bp for average insert size  
1.8kb. Library constructed by Ling Hong in the laboratory  
of Gerald M. Rubin (University of California, Berkeley)  
using ZAP-cDNA synthesis kit (Stratagene) and Superscript  
II RT (Life Technologies). Note: this is a NIH\_MGC  
Library..!"

BASE COUNT 155 a 245 c 224 g 118 t  
ORIGIN

alignment\_scores:  
Quality: 1049.00 Length: 201  
Ratio: 5.245 Gaps: 0  
Percent Similarity: 99.502 Percent Identity: 99.502

alignment\_block:  
US-09-525-867-1 x BG765790 ..

Align seg 1/1 to: BG765790 from: 1 to: 742

13 ArgGleLeuGlyLeuArgSerValGlyProAlaValGlnAlaArgG1 29  
|||||  
3 CGGATCTTGTGCTGGCTCCAGCTGGGCTGGCTGTGCAGGACGAGG 52  
|||||  
29 yValHlsGlnSerValAlaThrAspGlyProSerSerThrGlnProAlaL 46  
|||||  
53 TETCCATCAGCGTGGCCACCGATGGCCCAAGCAGCAGCAGCTGCC 102  
|||||  
46 euProLysAlaArgAlaValAlaProLysProSerSerArgGlyGluTyr 62  
|||||



```
564 TTTGGGGCCNCAAGAACTTTTGGGTGNCCTGNTGNGCCCGGAGGTGA 515
94 etHisMetalala.ProArgTyAspMetAspArg.PheGlyValValP 110
111 TGNANATNGCAGCAACCCGNNAGACATGNCCTGTTTGGGGNGNTTT 465
110 heArgAlaSerProArgGlnSerAspValMetIleValAlaGlyThrLeu 126
464 TTTGGGCCCGCCCGCCAGTCTGAGTATGATTTGGCNGGNANANTA 415
127 ThrAsnLysMetalalaProAlaLeuArgLysValTyAspLysMetProG1 143
414 ACCAAAAGAGGGCCCGCCAGNCTTGC.AAGGTTTAAGACACAGATGCCGA 366
143 uprArgTyValValSerMetGlySerCysAlaAsnGlyGlyTyT 160
365 GCCCGANAGTGTTCACANGGGAGTGN.GCCAACGGNGGGTAAN 317
160 yrHisTySerTySerValValArgGlyCysAspArgIleValProVal 176
316 ACCACTATTCCTAATNGTGTGAGGGGNTGCCACCGCATTTGCCCGTG 267
177 AspIleTyIleProGlyCysProProThrAlaGluAlaLeuTyG1 193
266 GAATTTAAATCCAGGGTCCACACNANGCCGAGGCCCTGTTTANGG 217
193 yIleLeuGlnLeuGlnArgLysIleLysArgGluArgArgLeuGlnIleT 210
216 CATCCTGCAGTGCAGAGGAATCAAGCGGAGCGAGGCTGCANATT 167
210 rPTyArgArg 213
166 GGTACCGCAGG 156
```

seq\_name: /cgn2\_6/ptodata/2/ina/6B\_COMB.seq:us-09-030-607-32

```
seq_documentation_block:
; Sequence 32, Application US/09030607
; Patent No. 6262245
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY OF PROSTATE CANCER AND METHODS FO
; NUMBER OF SEQUENCES: 224
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED AND BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/030.607
; FILING DATE: 25-FEB-1998
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Maki, David J.
; REGISTRATION NUMBER: 31,392
; REFERENCE/DOCKET NUMBER: 210121.427C3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 32:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 789 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
```

```
; MOLECULE TYPE: cDNA
US-09-030-607-32

alignment_scores:
  Quality: 411.00      Length: 189
  Ratio: 3.067        Gaps: 5
  Percent Similarity: 70.899  Percent Identity: 58.201

alignment_block:
US-09-525-867-1 x US-09-030-607-32/rev ..
Align seg 1/1 to reverse of: US-09-030-607-32 from: 1 to: 789
```

```
36 ThrAspGlyProSerSerThrGlnProAlaLeuProLysAlaArgAlaVa 52
696 ACCGGGGGGCCCGAAGGGGNCANCTTCCNNCCAAAGNAGAGCGGTG 647
52 lAlaProLysProSerSerArgGlyGluTyValValAlaLysLeuAspA 69
646 GGTTCCCAAAACCCAGNAACGGGGGGGAGGATTTGGGGGCCAAGGGG 597
69 sp.LeuValAsnTrpAlaArgArgSerSerLeuTrpPro..... 81
596 ACTTTGTTG.....AAACTGGGNCCTCCCGGAGNTTT 565
82 .....MetThrPheGlyLeuAlaCysCysAlaValGluMetM 94
564 TTTGGGGCCNCAAGAACTTTTGGGTGNCCTGNTGNGCCCGGAGGTGA 515
94 etHisMetalala.ProArgTyAspMetAspArg.PheGlyValValP 110
514 TGNANATNGCAGCAACCCGNNAGACATGNCCTGTTTGGGGNGNTTT 465
110 heArgAlaSerProArgGlnSerAspValMetIleValAlaGlyThrLeu 126
464 TTTGGGCCCGCCCGCCAGTCTGAGTATGATTTGGCNGGNANANTA 415
127 ThrAsnLysMetalalaProAlaLeuArgLysValTyAspLysMetProG1 143
414 ACCAAAAGAGGGCCCGCCAGNCTTGC.AAGGTTTAAGACACAGATGCCGA 366
143 uprArgTyValValSerMetGlySerCysAlaAsnGlyGlyTyT 160
365 GCCCGANAGTGTTCACANGGGAGTGN.GCCAACGGNGGGTAAN 317
160 yrHisTySerTySerValValArgGlyCysAspArgIleValProVal 176
316 ACCACTATTCCTAATNGTGTGAGGGGNTGCCACCGCATTTGCCCGTG 267
177 AspIleTyIleProGlyCysProProThrAlaGluAlaLeuTyG1 193
266 GAATTTAAATCCAGGGTCCACACNANGCCGAGGCCCTGTTTANGG 217
193 yIleLeuGlnLeuGlnArgLysIleLysArgGluArgArgLeuGlnIleT 210
216 CATCCTGCAGTGCAGAGGAATCAAGCGGAGCGGAGGCTGCANATT 167
210 rPTyArgArg 213
166 GGTACCGCAGG 156
```

seq\_name: /cgn2\_6/ptodata/2/ina/6B\_COMB.seq:us-09-103-840A-2

```
seq_documentation_block:
; Sequence 2, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
```

```

; TITLE OF INVENTION: TUBERCULOSIS
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 4403765
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; FEATURE:
; OTHER INFORMATION: CDC 1551
; OTHER INFORMATION: "n" bases at various positions throughout the sequence
; OTHER INFORMATION: represent a, t, c or g
US-09-103-840A-2

alignment_scores:
    Quality: 449.00      Length: 220
    Ratio: 3.054        Gaps: 6
Percent Similarity: 66.818      Percent Identity: 41.818

alignment_block:
US-09-525-867-1 x US-09-103-840A-2 ..
Align seg 1/1 to: US-09-103-840A-2 from: 1 to: 4403765

4 LeuSerAlaProGlyLeuArgGlyPheArgIle..... 14
3507435ATTCATGCTCAGCGGTTCGTGGCTACGCGTATGTGGCGCGCGGGG 3507484
15 .....LeuGlyLeuArgSerSerValGlyProAlaValGlnAlaArgG 29
3507485GCTGACGTGGGATTGAGTAGGGGTGGGA..... 3507515

29 LyValHisGlnSerValAlaThrAspGlyProSerSerThrGlnProAla 45
3507516.....CTGGAAGAACAG 3507527
46 LeuProLysAlaArgAlaValAlaProLysProSerSerArgGlyGluTy 62
3507528CTGCCC.....GCGGGGAT 3507541
62 rValValAlaLysLeuAspAspLeuValAsnTrpAlaArgSerSerL 79
3507542CCTGTGTCGACCGTCGAGAGGTGGCGGTATGTCGCAAAACTCC 3507591
79 euTrpProMetThrPheGlyLeuAlaCysCysAlaValGlnMetMetHis 95
3507592TGTGGCGGCAACATTCGATTGGCGTCTGTGGATCGAGATGCGG 3507641
96 MetAlaAlaProArgTyAspMetAspArgPheGlyValVal...PheAr 111
3507642ACGCGGGACCAAGTTTGACATTGCGCGGTTCGGATGGAACGTTCTC 3507691
111 gAlaSerProArgGlnSerAspValMetIleValAlaGlyThrLeuThrA 128
3507692GGCGCGCGGACGAGATCTGATGATCTGTCGCGCGCGGTGACGC 3507741
128 solysMetAlaProAlaLeuArgLysValTyAspGlnMetProGluPro 144
3507742AGAAGATGCGCGGTACTGCGCCAGATCTATGACAGATGCGGAGCCG 3507791
145 ArgTyValValSerMetGlySerCysAlaAsnGlyGlyTyTyTrpHi 161
3507792AAATGGGTTCTGGCCATGGTGTGTGCGCCCTCGTCAGGTGGATGTTCAA 3507841
161 sTySerTySerValValArgGlyCysAspArgIleValProValAsp 178
3507842C...AACTATGCATCGTCGAGGGCGTGATGTTGTCGCTCGACA 3507888
178 leTyTrpProGlyCysProProThrAlaGluAlaLeuLeuTyGlyIle 194

```

```

3507889TCTACTACCCGGGTGCCCGCGCGCGAGATGCTGCTGCACGCAATC 3507938
195 LeuGlnLeuGlnArgLysIleLys.....ArgG1 204
3507939CTGAAGCTGCAGAAAGATTCAGCAGATGCCATTAGGTATCAACCGGA 3507988
204 uArgArgLeu 207
3507989ACGCGCTATC 3507998
seq_name: /cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-103-840A-1
seq_documentation_block:
; Sequence 1, Application US/09103840A
; Patent NO. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 4411529
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; OTHER INFORMATION: H37Rv
US-09-103-840A-1

```

```

alignment_scores:
    Quality: 449.00      Length: 220
    Ratio: 3.054        Gaps: 6
Percent Similarity: 66.818      Percent Identity: 41.818

alignment_block:
US-09-525-867-1 x US-09-103-840A-1 ..
Align seg 1/1 to: US-09-103-840A-1 from: 1 to: 4411529

4 LeuSerAlaProGlyLeuArgGlyPheArgIle..... 14
3511999ATTCATGCTCAGCGGTTCGTGGCTACGCGTATGTGGCGCGCGGG 3512048
15 .....LeuGlyLeuArgSerSerValGlyProAlaValGlnAlaArgG 29
3512049GCTGACGTGGGATTGAGTAGGGGTGGGA..... 3512079
29 LyValHisGlnSerValAlaThrAspGlyProSerSerThrGlnProAla 45
3512080.....CTGGAAGAACAG 3512091
46 LeuProLysAlaArgAlaValAlaProLysProSerSerArgGlyGluTy 62
3512092CTGCCC.....GCGGGGAT 3512105
62 rValValAlaLysLeuAspAspLeuValAsnTrpAlaArgSerSerL 79
3512106CTGTGTCGACCGTCGAGAGGTGGCGGTATGTCGCAAAACTCC 3512155
79 euTrpProMetThrPheGlyLeuAlaCysCysAlaValGlnMetMetHis 95
3512156TGTGGCGGCAACATTCGATTGGCGTCTGTGGATCGAGATGCGG 3512205
96 MetAlaAlaProArgTyAspMetAspArgPheGlyValVal...PheAr 111
3512206ACCGCGGACCAAGTTTGACATTGCGCGGTTCGGATGGAACGTTCTC 3512255

```

```

111 gAlaSerProArgGlnSerAspValMetIleValAlaGlyThrLeuThra 128
|||||.....
3512256GGCCACGCGCGGAGAGATCGTGGCGGCGCGGTGAGCC 3512305

128 snLysMetAlaProAlaLeuArgLysValTyrAspGlnMetProGluPro 144
|||||.....
3512306AAGAGATGCGCGCGTACTGCGCGAGATCATGACAGATGGCGGAGCCG 3512355

145 ArgTyrValValSerMetGlySerCysAlaAsnGlyClyGlyTyrHI 161
|||||.....
3512356AAATGGGTCTGGCCATGGGTGTGGCGCTCGTACGTTGGGATGTCAA 3512405

161 sTyrSerTyrSerValValArgGlyCysAspArgIleValProValAspi 178
|||||.....
3512406C...AATATCGATCGTGCAGGCGGTGATCATGTGTTTCGGTGCACA 3512452

178 leTyrIleProGlyCysProProThrAlaGluAlaLeuTyrGlyIle 194
|||||.....
3512453TCTACCTACCGCGGTGCGCGCGCGCGGAGATGCTGTGCACGAATC 3512502

195 LeuGlnLeuGlnArgLysIleLys.....ArgGI 204
|||||.....
3512503CTGAAGCTGCACGAAGATTCACAGATGCCATTAGGTATCAACCGGA 3512552

204 uArgArgLeu 207
|||||...
3512553ACGCGCTATC 3512562

```

```
seq_name: /cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-801-344-7
```

```

seq_documentation_block:
; Sequence 7, Application US/08801344
; Patent No. 6087140
; GENERAL INFORMATION:
; APPLICANT: Cameron, Douglas C.
; APPLICANT: Shaw, Anita J.
; APPLICANT: Altaras, Nedim E.
; TITLE OF INVENTION: MICROBIAL PRODUCTION OF 1,2-PROPANEDIOL
; TITLE OF INVENTION: FROM SUGAR
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Demitt Ross & Stevens S.C.
; STREET: 8000 Excelsior Drive, Suite 401
; CITY: Madison
; STATE: WI
; COUNTRY: U.S.A.
; ZIP: 53717-1914
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/801.344
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Sara, Charles S.
; REFERENCE/DOCKET NUMBER: 09820.037
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 608-831-2100
; TELEFAX: 608-831-2106
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3524 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Pyridine nucleotide transhydrogenase gene

```

```
US-08-801-344-7
```

# alignment\_scores:

```

Quality: 101.00 Length: 218
Ratio: 0.971 Gaps: 11
Percent Similarity: 47.706 Percent Identity: 22.477

```

# alignment\_block:

```

US-09-525-867-1 x US-08-801-344-7 ..
Align seg 1/1 to: US-08-801-344-7 from: 1 to: 3524

      8 GlyLeuArgGlyPheArgIleLeuGlyLeuArgSerSerValGlyProAl 24
      |||||.....
1046 GGTGTTGAGTCTGGCGCGCATTTGGCGAGCAACAGTCTCGCGCGCAT 1095

      24 aValGlnAlaArgGlyValHisGlnSerValAlaThrAspGlyProSer 41
      |||||.....
1096 TGTG.....1099

      41 erThrGlnProAlaLeuProLysAlaArgAlaValAlaProLysProSer 57
      |||||.....
1100 .....CGTGCATTCGACACCGCGCGAA 1123

      58 Ser ArgGlyGluTyrValValAlaLysLeuAspAspLeuValAsnTrpA 74
      |||||.....
1124 GTGAAGACAAAGTTCAACTA.....TGGG 1149

      74 laArgArgSerSerLeuTrpProMetThrPheGlyLeuAlaCysCys... 89
      |||||.....
1150 CGCGAATTCCTCGAGCTGATTTTAAAGAGGAGGTGGCAGCGCGATG 1199

      90 AlaVal..GluMetMetHisMetAlaAlaProArgTyrAspMetAspArg 105
      |||||.....
1200 GCTATGCCAAAGTGTATCGGACGCGTTCATCAACGCGGAATGGAA... 1246

      106 PheGlyValValPheArgAlaSerProArgGlnSerAspValMetIleVa 122
      |||||.....
1247 .....CTCTTTGCGGCCAGGCAAAAGAGTGTGATATCATTTGTCAC 1287

      122 laLaGlyThrLeuThrAsnLysMetAlaProAlaLeu.....ArgLysV 137
      |||||.....
1288 CACCGCGCTTATCCAGGCAACACCGCGCGAAGCTAATTACCGTGAAA 1337

      137 alTyrAspGlnMetProGluProArgTyrValValSerMetGlySerCys 153
      |||||.....
1338 TGGTTGACTCCATGAAGCGGCGGAGTGTGATTGTGCACCTG...GCAGCC 1384

      154 AlaAsnGlyGlyTyrTyrHisTyrSerTyrSerValValArgGlyCy 170
      |||||.....
1385 CAAAACGGCGGC.....AACTG 1401

      170 sAspArgIleValProValAspIleTyr.....179
      |||||.....
1402 TGAATACACCGTCCCGGTGAAATCTTCATCAGCGAAATGGTGTCAAAG 1451

      180 .....IleProGlyCysProProThrAlaGluAlaLeu 190
      |||||.....
1452 TGATTGTTTATACCGATCTTCGCGGCGCTGTGCGACGAATCCTCACAG 1501

      191 LeuTyrGly.....IleLeuGlnLeuGlnArgLysIleLysArgGluAr 205
      |||||.....
1502 CTTTACGGCACAAACCTCGTTATCTGCTGAACACTGTTGTGCAAGAGAA 1551

      205 g 205
      :
      1552 A 1552

```

```
seq_name: /cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-498-599-7
```

```

seq_documentation_block:
; Sequence 7, Application US/09498599

```

```
; Patent No. 6303352
; GENERAL INFORMATION:
; APPLICANT: Cameron, Douglas C.
; APPLICANT: Shaw, Anita J.
; APPLICANT: Altaras, Nedim E.
; TITLE OF INVENTION: MICROBIAL PRODUCTION OF
; 1,2-PROPANEDIOL FROM SUGAR
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dewitt Ross & Stevens S.C.
; STREET: 8000 Excelsior Drive, Suite 401
; CITY: Madison
; STATE: WI
; COUNTRY: U.S.A.
; ZIP: 53717-1914
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/498,599
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Sara, Charles S.
; REFERENCE/DOCKET NUMBER: 09820.037
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 608-831-2100
; TELEFAX: 608-831-2106
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3524 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Pyridine nucleotide transhydrogenase gene
; US-09-498-599-7

alignment_scores:
  Quality: 101.00      Length: 218
  Ratio: 0.971         Gaps: 11
  Percent Similarity: 47.706   Percent Identity: 22.477

alignment_block:
  US-09-525-867-1 x US-09-498-599-7  ..
  Align seg 1/1 to: US-09-498-599-7 from: 1 to: 3524
      8 GlyLeuArgGlyPheArgIleLeuGlyLeuArgSerSerValGlyProAl 24
      1046 GTGTGTCAGGCTGCGCCCATTTGGCGAGCAACAGTCTCGCGCGCAT 1095
      24 aValGlnAlaArgGlyValHisGlnSerValAlaThrAspGlyProSers 41
      1096 TGTG.....
      41 erThrGlnProAlaLeuProLysAlaArgAlaValAlaProLysProSer 57
      1100 .....CGTGCATTGACACCGCCGGAA 1123
      58 Ser.ArgGlyGlyThrValValAlaLysLeuAspLeuValAsnTPA 74
      1124 GTGAAGACAACTTCAAGTA.....TGGG 1149
      74 laArgArgSerLeuTrpProMetThrPheGlyLeuAlaCysCys... 89
      1150 CGGGGAATCTCGAGCTGGATTTTAAAGAGGAGAGCTGCACGCGCGATG 1199
```

```
90 AlaVal..GluMetMetHisMetAlaAlaProArgTyrAspMetAspArg 105
   |||:..:..:..:..:..:..:..:..:..:..:..:..:..:..:..:..:
1200 GCTATGCCAAAGTGATGTCGGACGCGTTTCATCAAGCGGAATGGAA... 1246
106 PheGlyValValPheArgAlaSerProArgGlnSerAspValMetIleVa 122
   :..:..:..:..:..:..:..:..:..:..:..:..:..:..:..:..:
1247 .....CTCTTTCGCCGCCAGCAAGAGAGGTGCATATCATTTGCAC 1287
122 lAlaGlyThrLeuThrAsnLysMetAlaProAlaLeu.....ArgLysV 137
   :..:..:..:..:..:..:..:..:..:..:..:..:..:..:..:..:
1288 CACCGGCCTATTCCAGGCAACACCGCGCGGAAGCTAATTACCCGTGAA 1337
137 alTyrAspGlnMetProGluProArgTyrValValSerMetGlySerCys 153
   :..:..:..:..:..:..:..:..:..:..:~::~:~::~:~::~:~::~:
1338 TGGTTGACTCCATGAAGCGGCGAGTGTGATTGTGCACCTG...GCAGCC 1384
154 AlaAsnGlyGlyTyrTyrHisTyrSerTyrSerValValArgGlyCy 170
   |||:~::~:~::~:~::~:~::~:~::~:~::~:~::~:~::~:~::~:~::~:
1385 CAAACGCGCGC.....AACTG 1401
170 sAspArgIleValProValAspIleTyr..... 179
   |||:~::~:~::~:~::~:~::~:~::~:~::~:~::~:~::~:~::~:~::~:
1402 TGAATACACGTCGCGGTGAAATCTTCATACGGAAATGGTGTCAAAG 1451
180 .....IleProGlyCysProProThrAlaGluAlaLeu 190
1452 TGATTGGTTATACCGATCTCCGGGCGCTCGCGACGCAATCCTCACAG 1501
191 LeuTyrGly.....IleLeuGlnLeuGlnArgLysIleLysArgGluAr 205
   |||:~::~:~::~:~::~:~::~:~::~:~::~:~::~:~::~:~::~:~::~:
1502 CTTTAGCGCACAAACCTCGTTATCTGCTGAAACTGTTGTGCAAGAGAA 1551
205 g 205
1552 A 1552

seq_name: /cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-178-252-26

seq_documentation_block:
; Sequence 26, Application US/09178252
; Patent No..6218188
; GENERAL INFORMATION:
; APPLICANT: Cardineau, Guy A.
; APPLICANT: Steelman, Steven J.
; APPLICANT: Narva, Kenneth E.
; TITLE OF INVENTION: Plant-Optimized Genes Encoding Pesticidal Toxins
; FILE REFERENCE: MA-714XC2
; CURRENT APPLICATION NUMBER: US/09/178,252
; CURRENT FILING DATE: 1998-10-23
; EARLIER APPLICATION NUMBER: 60/065,215
; EARLIER FILING DATE: 1997-11-12
; EARLIER APPLICATION NUMBER: 60/076,445
; EARLIER FILING DATE: 1998-03-02
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: patentin ver. 2.0
; SEQ ID NO 26
; LENGTH: 1965
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic B.t. toxin gene
; US-09-178-252-26

alignment_scores:
  Quality: 86.50      Length: 200
  Ratio: 0.911         Gaps: 8
  Percent Similarity: 47.500   Percent Identity: 26.500

alignment_block:
  US-09-525-867-1 x US-09-178-252-26/rev  ..
```

Align seg 1/1 to reverse of: US-09-178-252-26 from: 1 to: 1965

```
4 LeuSerAlaProGlyLeuArgGlyPheArg.....lleLeuGlyLe 17
|||||: ||||| ||| |||
1191 CTGGCGTAGGAGGCTCTGTAGTGGTCTGTGGTCTGTGGACT 1142
|||||: ||||| ||| |||
17 uArgSerValGlyProAlaVal...GlnAlaArgGlyValHisGlns 33
|||||: ||||| ||| |||
1141 GCAGGGTGGCGGACAGTAGTCAGGAAGAGTGGTGGTGGGGCG 1092
|||||: ||||| ||| |||
33 erValAlaThrAspGlyProSerThrGlnProAlaLeuProLysAla 49
|||||: ||||| ||| |||
1091 GGTAGCGGTTCGGGAGATGTTGAGCGGTTGAGCGCTCGAAGAGTG 1042
|||||: ||||| ||| |||
50 ArgAlaValAlaProLysProSerSerArgGlyGlyValVal..... 64
|||||: ||||| ||| |||
1041 CGGCGCCTGTATGAAGCGGTCTCCAGTCGGAAGGTGTTTACGGGT 992
|||||: ||||| ||| |||
65 .....AlaLysLeuAspAspLeuValAsnTrpAlaArgArgSers 78
|||||: ||||| ||| |||
991 TGTTCGCCAGCGCGGAGATGCCCTGGT.....GGCGGCGG.... 952
|||||: ||||| ||| |||
78 erLeuTrpProMetThrPheGlyLeuAlaCysCysAlaValGluMetMet 94
|||||: ||||| ||| |||
951 .....GTTGTA 946

95 HisMetAlaAlaProArgTyAspMetAspArgPheGlyValValPheAr 111
|||||: ||||| ||| |||
945 CAC.....GATCGGTGGTGTAGATCTCGC 920

111 gAlaSerProArgGlnSerAspValMetIleValAlaGlyThrLeuThrA 128
|||||: ||||| ||| |||
919 CGGTGAGTGGGGTGGCGGCTGGGTACTGGCGACGCTGTGTAAC 870
|||||: ||||| ||| |||
128 snLysMetAlaProAlaLeuArgLysValTyAspGlnMetProGluPro 144
|||||: ||||| ||| |||
869 GGAAGAGCGGCACAGATCATCGCATAGGTCATCTCCCTGCGGAA 820

145 ArgTyValValSerMetGlySerCysAlaAsnGlyGlyGlyTyTyriH 161
|||||: ||||| ||| |||
819 GGGTGGTGTAGTGCCAGGACTC.....GGTGTGGTGGCGGCA 779

161 sTyrsSerTyrsValValArgGlyCysAspArgIleValProValAspI 178
|||||: ||||| ||| |||
778 CTTCTCCAGCGCGGTGTAGTCTCCAGCAGTGGTGGTGTAGAT. 730
|||||: ||||| ||| |||
178 leTyriIleProGlyCysProProThrAlaGluAlaLeuLeuTyriGly 194
|||||: ||||| ||| |||
729 .....GGGCTGGCTCTCTGGTGGGCTT 706
```

seq\_name: /cgn2\_6/prodata/2/ina/6A\_COMB.seq: US-09-036-987A-1

seq\_documentation\_block:  
; Sequence 1, Application US/09036987A  
; Patent No. 6143526

GENERAL INFORMATION:

APPLICANT: Baltz, Richard H.  
APPLICANT: Broughton, Mary C.  
APPLICANT: Crawford, Kathryn P.  
APPLICANT: Madduri, Krishnamurthy  
APPLICANT: Merlo, Donald J.  
APPLICANT: Treadway, Patti J.  
APPLICANT: Turner, Jan R.  
APPLICANT: Waldron, Clive

TITLE OF INVENTION: Biosynthetic Genes For Spinosyn Insecticide  
; NUMBER OF SEQUENCES: 39

CORRESPONDENCE ADDRESS:

ADDRESS: Dow AgroSciences LLC Patent Department  
STREET: 9330 Zionsville Road  
CITY: Indianapolis  
STATE: Indiana  
COUNTRY: USA

ZIP: 46268  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/036,987A  
FILING DATE: 09-MAR-1998  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Stuart, Donald R.  
REGISTRATION NUMBER: 28,479  
REFERENCE/DOCKET NUMBER: 50,608  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (317)337-4816  
TELEFAX: (317)337-4847  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 80161 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
US-09-036-987A-1

alignment\_scores:  
Quality: 86.00 Length: 216  
Ratio: 0.796 Gaps: 9  
Percent Similarity: 50.000 Percent Identity: 22.685

alignment\_block:

US-09-525-867-1 x US-09-036-987A-1 ..

Align seg 1/1 to: US-09-036-987A-1 from: 1 to: 80161

```
10 ArgGlyPheArgIleLeuGlyLeuArgSerSerValGlyProAlaValG 26
|||||: ||||| ||| |||
6480 CGTGATTCGCCGCCGCCGCGGATCCGAGCTCCAGCGCGGCTTCCACCAG 6529

26 nAla...ArgGlyValHisGlnSerValAlaThrAspGlyProSerSert 42
|||||: ||||| ||| |||
6530 TGCCACCGGTCTCTCGTCGAGCAGCGGCCGCGATGCGCGGATCGGCG 6579

42 hrGlnProAlaLeuProLysAlaArgAlaValAlaProLysProSerSer 58
|||||: ||||| ||| |||
6580 TCGATCCAGCCAGGAGTCGGG..... 6602

59 ArgGlyGluTyValValAlaAlaLysLeuAspLeuValAsnTrpAlaAr 75
|||||: ||||| ||| |||
6603 .....GATTCCTGTTCGCGTGGTAATC 6625

75 gArgSerSerLeuTrpProMetThrPheGlyLeuAlaCysCysAlaValG 92
|||||: ||||| ||| |||
6626 CGGTACGGATGCCAATATCCGGTCTTGGCATCCGACCATCTCAGTAG 6675

92 luMetMetHisMetAlaAlaProArgTyAspMetAspArg..... 105
|||||: ||||| ||| |||
6676 CGGTAGGCAAGTGTCTGCACAGAGTCACTCTGCGCGCCCTTCCGCC 6725

106 PheGlyValValPheArg..AlaSerProArgGlnSerAspValMetIleV 122
|||||: ||||| ||| |||
6726 AGCGCTCTGGTTTCCGGCTCTGCATCCAGGCGACGATCAGTC..... 6768

122 alAlaGlyThrLeuThrAsnLysMetAlaProAlaLeuArgLysValTy 138
|||||: ||||| ||| |||
6769 .....TTCGCGCTTGCCTTCAGGAGATGAGC 6795

139 AspGlnMetProGluProArgTyValValSerMetGlySerCysAlaAs 155
|||||: ||||| ||| |||
6796 GATGCCCGTGGCGAATCGCTTATGAGTCCCGGAGAGTGCTGTGT 6845
```

```

92 luMetWethHisMetAlaAlaProArgTyrAspMetAspArg..... 105
   :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
6676 CGGTAGCGCAAGTGTGCTACACGAAGTCACCTCTGGCGGCGCCTTCCCC 6725
106 PheGlyValValPheArg, AlaSerProArgInSerAspValMetIleV 122
   :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
6726 AGCGCTCTGTGTTTCCGGCTCTGCATGCGAGCGGACCATCAGTC..... 6768
122 alAlaGlyThrLeuThrAsnLysMetAlaProAlaLeuArgLysValTyr 138
   :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
6769 .....TTGCGGCTTGTCCTCAGGAGATGAGC 6795
139 AspGlnMetProGluProArgTyrValValSerMetGlySerCysAlaAs 155
   :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
6796 GATGCCCGTGGCGANTCGTWTATGACGTGCCCGGACGAGTGTCTGT 6845
155 nGlyGlyGlyTyrHisTyrSer...TyrSerValValArgGlyCys... 170
   :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
6846 CTCGGCGCCTTACACCTTCTCGCCTGTGTTGATGCGGTGCGGAGCATCA 6895
171 .....AspArgIleValProValAspIleTyr...Ile 180
   :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
6896 GGACACGCGGAGCAAGGAAAGCGCTCATTTGACTCAGAAATCTCGATCTA 6945
181 ProGlyCysProProThrAlaGluAlaLeuLeuTyrGlyIleLeu..... 195
   :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
6946 CCGGCGCACCCGACTCGGTAGAGGCCAGGCTAGCGGAGACGACCTGCTC 6995
196 .GlnLeuGlnArgLysIleLysArgGluArgGluGlnIleTyr 210
   :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
6996 GCGCTGTGCAAGATCGCTACCATCATCCTCGAAGGCTAAGATTGG 7041
seq_name: /cqn2_6/ptodata/2/ina/5B_COMB.seq:10S-08-834-655-3

```

```

seq_documentation_block:
; Sequence 3, Application US/08834655
; Patent No. 5968809
; GENERAL INFORMATION:
; APPLICANT: KNUITZON, DEBORAH
; APPLICANT: MURKERJI, PRADIP
; APPLICANT: HUANG, YUNG-SHENG
; APPLICANT: THURMOND, JENNIFER
; APPLICANT: CHAUDHARY, SUNITA
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR SYNTHESIS
; OF LONG CHAIN POLY-UNSATURATED FATTY ACIDS IN PLANTS
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: RAE-VENTER LAW GROUP, P. C.
; STREET: 260 SHERIDAN AVENUE, P.O. BOX 60039
; CITY: PALO ALTO
; STATE: CA
; COUNTRY: USA
; ZIP: 94306
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/834,655
; FILING DATE: 11-APR-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: RAE-VENTER, BARBARA
; REGISTRATION NUMBER: 32,750
; REFERENCE/DOCKET NUMBER: CGNE.124.000S
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 328-4400
; TELEFAX: (650) 328-4477
; TELEX: N/A
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:

```



```
; LENGTH: 1488 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-834-655-3

alignment_scores:
  Quality: 84.50      Length: 281
  Ratio: 0.741       Gaps: 16
  Percent Similarity: 40.569  Percent Identity: 20.996

alignment_block:
US-09-525-867-1 x US-08-834-655-3

Align seg 1/1 to: US-08-834-655-3 from: 1 to: 1488

25 ValGlnAlaArgGlyValHisGlnSerValAlaThrAspGlyProSerSe 41
196 ATCGATGCGGGTTTGACCCAGCGTCATATCAGCACCTCGGCCCAACTC 245
41 rThGlnProAlaLeuProLysAlaArgAlaValAlaProLysProSer 57
246 GGCCAAGCTGCTTCGAGCGCAACTACCAGCTCCCGAGTTCACCATCA 295
58 SerArgGlyGlu..... 61
296 AGGAGATCCGAGAGTGCATCCCTGCCACTGCTTGGAGCGCTCCGGTCTC 345
62 TyrValValAlaLysLeuAspLeu.....ValA 72
346 CGTGTCTCTGCGAGTGGCATCGATCGATCGACTGGCGTGGCTCTGTGT 395
72 snTtpAlaArgArg.....SerSerLeu 79
396 CTGGCTGCGACCCAGATCGACAGTTTGAGATCCCTGATCCGCTATT 445
80 TrpProMetThrPheGlyLeuAlaCysAlaValGluMetMethisMe 96
446 TGGCTGCGCTGTTTACTGGATCATGACGAGGTA.....TT 480
96 talaAlaProArgTyrAspMetAspArgPheGlyValValPheArgAla 113
481 GTCTGACCGGTGCTGGGTGCTGGCTCAGAGTGGTGCATCATGCTCTT 530
113 er.....ProArgGlnSerAspValMetileValAlaGlyThrLeuThr 127
531 CTGACCTCCAAAGACCTCAACACACAGTGTGT...GGATCTTGCAC 577
128 AsnLysMetAlaPro.....AlaLeuArgLys. 136
578 CGATGCTCTGGTCCCTTACCCTGAGATCTCGCACTCGAAGCAC 627
137 .....ValTyrAspGlnMetProglu..... 143
628 CACAAAGCCACTGCGCATATGACCAAGACAGTGTGTGTCGCCAAGAC 677
144 .....ProArgTyrValValSerMetGlySer.....CysAlaAs 155
678 CCGTCCCAAGTTGGCTGGCTGCCAAGGAGAGAGTGTGCTGCGCGTTC 727
155 nGlyGlyGlyTyrThrHisTyrSerTyrSerValValArgGly..... 169
728 AGGAGAGGAGCAT.....GTCCGTTGCACCTGGAT 756
170 .....CysAsp..... 171
757 GAGGAGGCTCCCATGTGTGACTTTTCTGGATGCTGATCCAGTCTTGT 806
172 .....ArgileValPr 175
807 CGGATGCGCCGCGTACTGTATTATGAACGCTCTGCCAAGAGCATCGGC 856
```

```
175 oValAspileTyrIlePro..... 181
857 GCTGACCTCGCACTTCCACACGTAAGTCTGCGCCATCTTTGAGCCCCCAAC 906
182 .....GlyCysProProThr 186
907 TTTTTCGACATTATTATCTCGGACCTCGGTGTGTGGCTGCCCTCGGTGC 956
187 AlaGluAlaLeuTyrGlyIleLeuGlnLeuGlnArg 199
957 CCTGATCTATGCTCCATGCGAGTGTGCTCTTGGACCGT 995

seq name: /cqn2_6/ptodata/2/ina/6A_COMB.seq:US-08-834-033A-3

seq_documentation_block:
; Sequence 3, Application US/08834033A
; Patent No. 6075183
; GENERAL INFORMATION:
; APPLICANT: KNUDSON, DEBORAH
; APPLICANT: MUKERJI, PRADIP
; APPLICANT: HUANG, YUNG-SHENG
; APPLICANT: THURMOND, JENNIFER
; APPLICANT: CHAUDHARY, SUNITA
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR SYNTHESIS
; OF LONG CHAIN POLY-UNSATURATED FATTY ACIDS IN PLANTS
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LIMBACH AND LIMBACH, L.L.P.
; STREET: 2001 FERRY BUILDING
; CITY: SAN FRANCISCO
; STATE: CA
; COUNTRY: USA
; ZIP: 94111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WORDPERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/834,033A
; FILING DATE: 11-APR-1997
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: WARD, MICHAEL R.
; REGISTRATION NUMBER: 38,651
; REFERENCE/DOCKET NUMBER: CGAB-300.USA
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 433-4150
; TELEFAX: (415) 433-8716
; TELEX: N/A
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1488 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-834-033A-3

alignment_scores:
  Quality: 84.50      Length: 281
  Ratio: 0.741       Gaps: 16
  Percent Similarity: 40.569  Percent Identity: 20.996

alignment_block:
US-09-525-867-1 x US-08-834-033A-3

Align seg 1/1 to: US-08-834-033A-3 from: 1 to: 1488

25 ValGlnAlaArgGlyValHisGlnSerValAlaThrAspGlyProSerSe 41
.....
```

```
196 ATCGATGCGGTTTGACCGAGGTGATATCAGCACTCGGCCCAAACTC 245
41 rThrGlnProAlaLeuProLys..AlaArgAlaValAlaProLysProSer 57
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
246 GGCAAGCGTCTGCTTGGAGCGCAACTACCAGCTCCCGGAGTTCAACATCA 295
58 SerArgGlyGlu..... 61
   ||||| |||||
296 AGGAGATCGGAGAGTGCATCCCTGCCACTGCTTTGAGCGGTCCGGTCTC 345
62 .TyrValValAlaLysLeuAspLeu.....Vala 72
   ||||| ||||| ||||| ||||| ||||| ||||| |||||
346 CGTGGTCTCTGCCACCTGCTCCATCGATCTGACTTGGCGTCCGCTCTTGT 395
72 snTrpAlaArg.....SerSerLeu 79
   ||||| |||||
396 CTTGGTGGCGAGCGAGATCGACAAGTTTGAGAATCCCTTGATCCGCTATT 445
80 TrpProMetThrPheGlyLeuAlaCysCysAlaValGluMetMetHisMe 96
   ||||| ||||| ||||| ||||| ||||| ||||| |||||
446 TGGCTGGCTGTTTACTGGATCATGACGGTA.....TT 480
96 tAlaAlaProArgTyrAspMetAspArgPheGlyValValPheArgAlas 113
   ||||| ||||| ||||| ||||| ||||| ||||| |||||
481 GTCTGACCGGTGCTCTGGTGGTGGTCTGACGAGTGTGTATCATCATCTCT 530
113 er.....ProArgGlnSerAspValMetIleValAlaGlyThrLeuThr 127
   ||||| ||||| ||||| ||||| ||||| ||||| |||||
531 CTGACCTCAAGACCTTCAACACACAGTTGGT....GGATCTTGCACCT 577
128 .AsnLysMetAlaPro.....AlaLeuArgLys. 136
   ||||| ||||| ||||| ||||| ||||| ||||| |||||
578 CGATGCTCTGGTCCCTACCACCTCTGAGAGATCTCGCACTCGAAGCAC 627
137 .....ValTyrAspGlnMetProGlu..... 143
628 CACAAGGCCACTGGCCATATGACCAAGGACAGGTCTTTGTGCCCAAGAC 677
144 .....ProArgTyrValValSerMetGlySer.....CysAlaAs 155
   ||||| ||||| ||||| ||||| ||||| ||||| |||||
678 CCGCTCCAGGTGGTGTGCTCCCAAGGAGAACGCTGCTGCTGCGCTTC 727
155 nGlyGlyTyrTyrHisTyrSerTyrSerValValArgGly..... 169
   ||||| ||||| ||||| ||||| ||||| ||||| |||||
728 AGGAGAGGACAT.....GTCCGTGCACCTGGAT 756
170 .....CysAsp..... 171
757 GAGGAGCTCCCATTTGACTTTGTTCTGGATGGTGATCCAGTTCTTCTT 806
172 .....ArgIleValPr 175
   ||||| ||||| ||||| ||||| ||||| ||||| |||||
807 CGGATGCCCGCGTACCTGATTATGAACCCCTGTGCCAAGACTACGGCC 856
175 oValAspIleTyrIlePro..... 181
857 GCTGGACCTCGCACTTCCACAGCTACTGCCCATCTTTGAGCCCGGCAAC 906
182 .....GlyCysProProThr 186
   ||||| ||||| ||||| ||||| ||||| ||||| |||||
907 TTTTTCGACATTATTATTCGGACCTCGGTGTGTTGGCTGCGCTCGGTGC 956
187 AlaGluAlaLeuTyrGlyIleLeuGlnLeuArg 199
   ||||| ||||| ||||| ||||| ||||| ||||| |||||
957 CCGTATCTATGCTCATCATGCTTGTGCTCTTGCCTTGCCTT 995
```

seq\_name: /cgn2\_6/ptodata/2/ina/6A\_COMB.seq:US-09-363-574-3

seq\_documentation\_block:

; Sequence 3, Application US/09363574  
; Patent No. 6136574  
; GENERAL INFORMATION:  
; APPLICANT: KNUTZON, DEBORAH

```
; APPLICANT: MURKERJI, PRADIP
; APPLICANT: HUANG, YUNG-SHENG
; APPLICANT: THURMOND, JENNIFER
; APPLICANT: CHAUDHARY, SUNITA
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR SYNTHESIS
; TITLE OF INVENTION: OF LONG CHAIN POLY-UNSATURATED FATTY ACIDS IN PLANTS
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESS: LIMBACH AND LIMBACH L.L.P.
; STREET: 2001 FERRY BUILDING
; CITY: SAN FRANCISCO
; STATE: CA
; COUNTRY: USA
; ZIP: 94111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/363,574
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: WARD, MICHAEL R.
; REGISTRATION NUMBER: 38,651
; REFERENCE/DOCKET NUMBER: CGAB-202 USA
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 433-4150
; TELEFAX: (415) 433-8716
; TELEX: N/A
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1488 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-09-363-574-3
```

alignment\_scores:  
Quality: 84.50 Length: 281  
Ratio: 0.741 Gaps: 16  
Percent Similarity: 40.569 Percent Identity: 20.996

alignment\_block:

US-09-525-867-1 x US-09-363-574-3 ..

Align seg 1/1 to: US-09-363-574-3 from: 1 to: 1488

```
25 ValGlnAlaArgGlyValHisGlnSerValAlaThrAspGlyProSerSe 41
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
196 ATCGATGCGGTTTGACCGAGGTGATATCAGCACTCGGCCCAAACTC 245
41 rThrGlnProAlaLeuProLys..AlaArgAlaValAlaProLysProSer 57
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
246 GGCAAGCGTCTGCTTGGAGCGCAACTACCAGCTCCCGGAGTTCAACATCA 295
58 SerArgGlyGlu..... 61
   ||||| |||||
296 AGGAGATCGGAGAGTGCATCCCTGCCACTGCTTTGAGCGGTCCGGTCTC 345
62 .TyrValValAlaLysLeuAspLeu.....Vala 72
   ||||| ||||| ||||| ||||| ||||| ||||| |||||
346 CGTGGTCTCTGCCACCTGCTCCATCGATCTGACTTGGCGTCCGCTCTTGT 395
72 snTrpAlaArg.....SerSerLeu 79
   ||||| |||||
396 CTTGGTGGCGAGCGAGATCGACAAGTTTGAGAATCCCTTGATCCGCTATT 445
80 TrpProMetThrPheGlyLeuAlaCysCysAlaValGluMetMetHisMe 96
   ||||| ||||| ||||| ||||| ||||| ||||| |||||
```



```

150 tclY.....SerCysAlaasnG 156
   :::::
1318 GTCTCGTCGGGTACGTCCTCCGCCGACACCGGTGCACGTGTCTCGGA 1367
   :::::
156 lYlGlylTyTy-HisTyrSerValValArglYcysAspArg 172
   |||||
1368 CCGCGGCG.....GGGTCCGGGAAG 1387
   |||||
173 lIeVal.....ProValAspIleTyrII 180
   :::::
1388 TGCTGCTCAACATCGTCTCCACGGGATCAAGTACAATCGCTCGGTGT 1437
   :::::
180 eProglYcysProProThrAlaGluAlaLeuLeuTyrGlyIleLeuGlnL 197
   |||||
1438 GTCGTCGACCCCAACAGGATCAGGGCGTCTCGTCGGGTACAGCAG 1487
   |||||
197 euGlnArglYsIlelYsArgGluArg 206
   :::::
1498 GGCCCCGGACTACCGACGAGCGCGCAGGA 1516
   :::::

```

seq\_name: /cgn2\_6/ptodata/2/ina/5A\_COMB.seq:US-08-577-184-5

```

seq_documentation_block:
: Sequence 5, Application US/08577184
: Patent No. 5602014
: GENERAL INFORMATION:
: APPLICANT: MIZUMURA, YURIE
: APPLICANT: YU, FUJIO
: TITLE OF INVENTION: A REGULATORY FACTOR FOR EXPRESSION OF
: TITLE OF INVENTION: NITRILASE GENE AND A GENE THEREOF
: NUMBER OF SEQUENCES: 7
: CORRESPONDENCE ADDRESS:
: ADDRESSER: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
: ADDRESSER: P.C.
: STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
: CITY: ARLINGTON
: STATE: VA
: COUNTRY: USA
: ZIP: 22202
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/577,184
: FILING DATE: 22-DEC-1995
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: JP 6-337652
: FILING DATE: 28-DEC-1994
: ATTORNEY/AGENT INFORMATION:
: NAME: OBLON, NORMAN F.
: REGISTRATION NUMBER: 24,618
: REFERENCE/DOCKET NUMBER: 2941-004-0
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 703-413-3000
: TELEFAX: 703-413-2220
: INFORMATION FOR SEQ ID NO: 5:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 2336 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: double
: TOPOLOGY: linear
: MOLECULE TYPE: DNA (genomic)
: ORIGINAL SOURCE:
: ORGANISM: RHODOCOCCUS ERYTHROPOLIS
: STRAIN: SK92
:
: US-08-577-184-5

```

[illegible]

```
seq_name: /cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-226-012-1
seq_documentation_block:
; Sequence 1, Application US/09226012
```

```

; Patent No. 6207383
; GENERAL INFORMATION:
; APPLICANT: Keating, Mark T.
; TITLE OF INVENTION: MUTATIONS IN AND GENOMIC STRUCTURE OF HERG - A LONG QT
; FILE OF INVENTION: SYNDROME GENE
; FILE REFERENCE: 2323-136
; CURRENT APPLICATION NUMBER: US/09/226,012
; CURRENT FILING DATE: 1999-01-06
; EARLIER APPLICATION NUMBER: 09/122,847
; EARLIER FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 116
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 3480
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(3477)
US-09-226-012-1

```

```

alignment_scores:
  Quality: 80.50      Length: 147
  Ratio: 1.134      Gaps: 6
  Percent Similarity: 48.299      Percent Identity: 27.211

alignment_block:
US-09-525-867-1 x US-09-226-012-1/rev ..

Align seg 1/1 to reverse of: US-09-226-012-1 from: 1 to: 3480

41 SerThrGlnProAlaLeuProLysAlaAlaValAlaProLysPro.. 56
||| ||||| ||| ||||| ||||| ||||| ||||| |||||
2363 TCGATGGACCCCGGAGATGAAGTACAGCGCGGTGAGCAGCTCCCGCAGC 2314

57 .....SerSerArgGlyGluTyrValValAlaLysLeuAspA 69
||| ||| ||| |||||
2313 ATGCACCACTGTGTCCCTCGCGGTGCATGTGTGTCTTG.....A 2273

69 spleuValAsnTrpAlaArgSerSerLeuTrpProMetThrPheGly 85
||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2272 ACTTCATGCCAGGCCCGAGAGCGCCCTTGTTGCCCTCGGAGCGGT 2223

86 LeuAlaCysCysAlaValGluMetMet.....HisMetAlaAlaPr 99
||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| |||
2222 TTGCAGTGTGCAGCAGTGCAGCGGTTCAGTGCAGGAGATGTCAGCGTG 2173

99 OArgTyrAspMetAspArgPheGlyValValPheArgAlaSerProArg 116
||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2172 CAGGCACTCAGGAAGCCCTTCAGCAGCGGTTCATGTCGATGCCGTGG 2123

116 InSerAspValMetIleValAlaGlyThrLeuThrAsnLysMetAlaPro 132
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
2122 TGTAGGACCGAGC....GTGCTGGAAGTACTCTCGAGGCGCTGGCGCAG 2077

133 AlaLeuArgLysValTyr.....AspGlnMetProGluProArgTyrVa 147
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
2076 GGGATTGGGATCTGTGTGAGAGCGGATGAACCTCCCGCAGCCGAGCATCT 2027

147 lValSerMetGlySerCysAlaAsnGlyGlyTyrTyrHisTyrSert 164
;
2026 G..... 2026

164 yrservValValArgGlyCysAspArgIleValProValAsp 177
||| ||||| ||||| ||||| ||||| ||||| |||||
2025 ..TGTTGGTAGCGGCTGTGCCCGAGTACAGCCGCTGGAT 1987

```

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 12, 2002, 19:53:05 ; Search time 15.62 Seconds  
(without alignments)  
306.863 Million cell updates/sec

Title: US-09-525-867-1  
Perfect score: 1118  
Sequence: 1 MAVLSAPGLRGFRILGRSS.....TLQLQKIKRRRLQIWR 213

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 212252 seqs, 22503292 residues  
Total number of hits satisfying chosen parameters: 212252

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued\_Patents\_AA:\*  
1: /cgn2\_6/ptodata/2/iaa/5A.COMB.pep.\*  
2: /cgn2\_6/ptodata/2/iaa/5B.COMB.pep.\*  
3: /cgn2\_6/ptodata/2/iaa/6A.COMB.pep.\*  
4: /cgn2\_6/ptodata/2/iaa/6B.COMB.pep.\*  
5: /cgn2\_6/ptodata/2/iaa/PTUS.COMB.pep.\*  
6: /cgn2\_6/ptodata/2/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	718	64.2	182	3	US-09-261-471-5
2	88.5	7.9	510	3	US-08-801-344-8
3	88.5	7.9	510	4	US-09-498-599-8
4	70.5	6.3	319	4	US-08-836-075A-42
5	70.5	6.3	425	1	US-08-615-170-17
6	70.5	6.3	432	1	US-08-615-170-3
7	70.5	6.3	445	1	US-08-615-170-5
8	70.5	6.3	497	3	US-08-258-287B-37
9	70.5	6.3	497	3	US-08-368-704C-37
10	70	6.3	4302	3	US-08-658-136-5
11	69.5	6.2	2955	2	US-08-443-260-3
12	69.5	6.2	2955	3	US-08-442-805A-3
13	69.5	6.2	2955	3	US-08-443-900A-3
14	69.5	6.2	2955	4	US-08-444-818-124
15	69.5	6.2	2955	4	US-08-249-843-3
16	69.5	6.2	3011	2	US-08-833-678A-6
17	69.5	6.2	3011	4	US-08-444-818-177
18	69.5	6.2	3011	4	US-08-529-169A-6
19	69	6.2	457	4	US-09-124-238A-10
20	68	6.1	310	5	PCT-US95-04801-6
21	68	6.1	395	1	US-08-318-947A-18
22	68	6.1	395	2	US-08-795-303-18
23	68	6.1	748	2	US-08-997-080-154
24	68	6.1	748	2	US-08-937-362-154
25	68	6.1	748	4	US-08-995-855-154
26	68	6.1	888	4	US-09-268-140-4
27	68	6.1	993	4	US-09-060-410-4

28	67.5	6.0	2995	4	US-08-444-818-138	Sequence 138, App
29	67	6.0	3025	6	5223423-3	Patent No. 5223423
30	67	6.0	4303	2	US-08-460-751-2	Sequence 2, Appl
31	66.5	5.9	665	2	US-08-846-762-17	Sequence 17, Appl
32	66.5	5.9	665	2	US-08-846-762-94	Sequence 94, Appl
33	66.5	5.9	967	1	US-08-188-281B-13	Sequence 13, Appl
34	66.5	5.9	967	5	PCT-US94-07280-13	Sequence 13, Appl
35	66.5	5.9	967	5	PCT-US95-01087-13	Sequence 13, Appl
36	66.5	5.9	1648	1	US-08-188-281B-12	Sequence 12, Appl
37	66.5	5.9	1648	5	PCT-US94-07280-12	Sequence 12, Appl
38	66.5	5.9	1648	5	PCT-US95-01087-12	Sequence 12, Appl
39	66.5	5.9	3011	1	US-08-188-281B-1	Sequence 1, Appl
40	66.5	5.9	3011	1	US-08-453-552-1	Sequence 1, Appl
41	66.5	5.9	3011	2	US-08-710-837-1	Sequence 1, Appl
42	66.5	5.9	3011	3	US-08-811-566-20	Sequence 20, Appl
43	66.5	5.9	3011	4	US-09-014-416-1	Sequence 1, Appl
44	66.5	5.9	3011	5	PCT-US91-02225-10	Sequence 10, Appl
45	66.5	5.9	3011	5	PCT-US93-00907-1	Sequence 1, Appl

ALIGNMENTS

RESULT 1  
US-09-261-471-5  
; Sequence 5, Application US/09261471  
; Patent No. 6025123  
; GENERAL INFORMATION:  
; APPLICANT: Bandman, Olga  
; APPLICANT: Lal, Preeti  
; APPLICANT: Corley, Neil C.  
; TITLE OF INVENTION: HUMAN SUCCINYL-COA SYNTHETASE HOLOENZYME  
; NUMBER OF SEQUENCES: 6  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.  
; STREET: 3174 PORTER DRIVE  
; CITY: PALO ALTO  
; STATE: CALIFORNIA  
; COUNTRY: USA  
; ZIP: 94304  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/261,471  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 09/099,677  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: CERRONE, MICHAEL C.  
; REGISTRATION NUMBER: 39,132  
; REFERENCE/DOCKET NUMBER: PF-0545 US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (650) 855-0555  
; TELEFAX: (650) 845-4166  
; INFORMATION FOR SEQ ID NO: 5:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 182 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; IMMEDIATE SOURCE:  
; LIBRARY: GenBank  
; CLONE: g2258365  
US-09-261-471-5

Query Match 64.2%; Score 718; DB 3; Length 182;  
Best Local Similarity 81.4%; Pred. No. 2.2e-76;

Matches	127;	Conservative	18;	Mismatches	11;	Indels	0;	Gaps	0;
---------	------	--------------	-----	------------	-----	--------	----	------	----

  

QY	58	SRGEYVAKDDLNNVARRSLWPMTFGLACCAVEMHMAAPRYMDRFGVVFRA	SPROS	117
Db	27	NKTEFIVSKMDLNNVARKSLWPMTFGLACCAVEMHMSAASRYDLDRFGIIFR	SPROS	86
QY	118	DYMIYAGILTNKMAPALRKVYDQMPERVYVVMGSCANGGYYHYYSYVVRGCDR	IVPVD	177
Db	87	DYMIYAGILTNKMAPALRKVYDQMPERVYVVMGSCANGGYYHYYSYVVRGCDR	IVPVD	146
QY	178	IYPGCPPTAEALLYGILQQRKIKRRLRIQIYWR	213	
Db	147	IYPGCPPTAEALLYGILQQRKIKRRLTIWLQK	182	

  

```

RESULT      2
US-08-801-344-8
; Sequence 8, Application US/08801344
; Patent No. 6087140
; GENERAL INFORMATION:
; APPLICANT: Cameron, Douglas C.
; APPLICANT: Shaw, Anita J.
; APPLICANT: Altaras, Nedim E.
; TITLE OF INVENTION: MICROBIAL PRODUCTION OF 1,2-PROPANEDIOL
; TITLE OF INVENTION: FROM SUGAR
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dewitt Ross & Stevens S.C.
; STREET: 8000 Excelsior Drive, Suite 401
; CITY: Madison
; STATE: WI
; COUNTRY: U.S.A.
; ZIP: 53717-1914
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/801.344
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Sara, Charles S.
; REFERENCE/DOCKET NUMBER: 09820.037
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 608-831-2100
; TELEFAX: 608-831-2106
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 510 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
; ORGANISM: Pyridine nucleotide transhydrogenase, subunit
; ORGANISM: A
; US-08-801-344-8

```

  

Query Match	7.9%;	Score 88.5;	DB 3;	Length 510;
Best Local Similarity	23.7%;	Pred. No. 0.079;		
Matches	27;	Conservative	23;	Mismatches 35;
Indels	29;	Gaps	5;	

  

QY	109	VFRASPROSDVMIVAGTILTNKMAPAL--RKYVDQMPERVYVVMGSCANGGYYHYYSYV	166
Db	241	LFAAQAKEVDILVTALIPGKPAKPLITREWVDSMKAGSIVIDL--AAQNGG-----	290
QY	167	VRGCDRIVPVDIY-----IPGCPPTAEALLYG--ILQQRKIKRER	205

;; TITLE OF INVENTION: NEW SEQUENCES OF HEPATITIS C VIRUS GENOTYPES  
;; TITLE OF INVENTION: AGENTS AND THEIR USE AS PROPHYLACTIC, THERAPEUTIC AND DIAGNOSTIC  
;; NUMBER OF SEQUENCES: 207  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: ARNOLD, WHITE & DURKEE  
;; STREET: P.O. BOX 4433  
;; CITY: HOUSTON  
;; STATE: TEXAS  
;; COUNTRY: USA  
;; ZIP: 77210-4433  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: Microsoft Word 6.0 / ASCII text output  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/836,075A  
;; FILING DATE: 21 Apr 1997  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: PCT/EP95/04155  
;; FILING DATE: 23 Oct 1995  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: EP 94870166.9  
;; FILING DATE: 21 Oct 1994  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: EP 95870076.7  
;; FILING DATE: 28 Jun 1995  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: KAMMERER, PATRICIA A.  
;; REGISTRATION NUMBER: 29,775  
;; REFERENCE/DOCKET NUMBER: INNS:004  
;; INFORMATION FOR SEQ ID NO: 42:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 319 amino acids  
;; TYPE: amino acid  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: peptide  
US-08-836-075A-42

Query Match 6.3%; Score 70.5; DB 4; Length 319;  
Best Local Similarity 19.5%; Pred. No. 5.1;  
Matches 45; Conservative 40; Mismatches 75; Indels 71; Gaps 14;  
Qy 7 PGLRGFRILG---LRSSVGPVAVQARGVHQSVATDGPSTQPALPKARAVAPKPSRGYEV 63  
Db 25 PG--GGQIVGGVILLPRGRRLGVRATRKTSERSQPRGRQPKYR---QPTGR---- 74  
Qy 64 VAKLDDLNNARRSSLNPMTEGLACCAVEMHMAAPRYDMDFG-----VVVF 110  
Db 75 -----SMQPGGYWPL-YGNEGCGWAGW-LLSPRGSRPNWGNDRRRSRNLGKVID 124  
Qy 111 RASPRQSDVM---IVAGTL---TNKMAPALRKVPD-----QMPD---PRVYVSMGSC 153  
Db 125 TLTXLADLMGYIVLGGPLGGVAAALAHGVRALEGVNVATGNLPGCSFSLALLSC 184  
Qy 154 -----ANGGGYHY-----SYSVWRCGRIVPVDIYIPGCPPTAEA 189  
Db 185 LTPATAIQVKNASGIYHLTNDCSNNSIVEAETMI---LHLPGCVPCIKA 232

RESULT 5  
US-08-615-170-17  
; Sequence 17, Application US/08615170  
; Patent No. 5776776  
; GENERAL INFORMATION:  
; APPLICANT: ORDAHL, Charles P.  
; APPLICANT: AZAKIE, Anthony  
; APPLICANT: MAR, Janet H.  
; APPLICANT: FARRANCE, Iain K.G.  
; APPLICANT: HALL, Deborah E.  
; APPLICANT: STEWART, Alexandre F.R.

APPLICANT: LARKIN, Sarah B.  
TITLE OF INVENTION: DTEF-1 ISOFORMS AND USES THEREOF  
NUMBER OF SEQUENCES: 32  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Townsend and Townsend Kourile and Crew  
STREET: Stewart Street Tower, One Market Plaza  
CITY: San Francisco  
STATE: California  
COUNTRY: US  
ZIP: 94105-1493  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/615,170  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/US95/01526  
FILING DATE: 06-FEB-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/191,493  
FILING DATE: 04-FEB-1994  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Heslin, James M.  
REGISTRATION NUMBER: 29,541  
REFERENCE/DOCKET NUMBER: 2307U-053120  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 326-2400  
TELEFAX: (415) 326-2422  
INFORMATION FOR SEQ ID NO: 17:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 425 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-615-170-17  
Query Match 6.3%; Score 70.5; DB 1; Length 425;  
Best Local Similarity 23.3%; Pred. No. 7.8;  
Matches 38; Conservative 20; Mismatches 62; Indels 43; Gaps 7;  
Qy 3 VLSAPGLRGFRIL-GLRSSVGPVAVQ-----RGVHQSV-----TDGPSSTQPALPKAR 50  
Db 130 IISATAPHSKMLPGLPRSAYPVSGVWQALPGQAGSQDVKPFQQPYALQPSLPLPG 189  
Qy 51 AVAP-----KPSSRGYVVAKLDDLNNARRSSLNPMTEGLACCAVEMHMAAPRYD 102  
Db 190 FDSPTGLPPSSPTPAWGRRVAS-----SKLWLEFS-----AFLEQQQODDTYN 234  
Qy 103 MDREGVYFRASPRQSDVMIVAGTLTNKMAPALRKVYDQMEPR 145  
Db 235 KHLEFVHIGQSNPSYSDPYL-----EAVDIRIQYDKPEKK 269

RESULT 6  
US-08-615-170-3  
; Sequence 3, Application US/08615170  
; Patent No. 5776776  
; GENERAL INFORMATION:  
; APPLICANT: ORDAHL, Charles P.  
; APPLICANT: AZAKIE, Anthony  
; APPLICANT: MAR, Janet H.  
; APPLICANT: FARRANCE, Iain K.G.  
; APPLICANT: HALL, Deborah E.  
; APPLICANT: STEWART, Alexandre F.R.  
; APPLICANT: LARKIN, Sarah B.



;; TITLE OF INVENTION: DTEP-1 ISOFORMS AND USES THEREOF  
;; NUMBER OF SEQUENCES: 32  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: Townsend and Townsend Kourie and Crew  
;; STREET: Stewart Street Tower, One Market Plaza  
;; CITY: San Francisco  
;; STATE: California  
;; COUNTRY: US  
;; ZIP: 94105-1493  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: PatentIn Release #1.0, Version #1.25  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/615,170  
;; FILING DATE:  
;; CLASSIFICATION: 435  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: PCT/US95/01526  
;; FILING DATE: 06-FEB-1995  
;; CLASSIFICATION: 435  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 08/191,493  
;; FILING DATE: 04-FEB-1994  
;; CLASSIFICATION: 435  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Heslin, James M.  
;; REGISTRATION NUMBER: 29,541  
;; REFERENCE/DOCKET NUMBER: 2307U-053120  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (415) 326-2400  
;; TELEFAX: (415) 326-2422  
;; INFORMATION FOR SEQ ID NO: 3:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 432 amino acids  
;; TYPE: amino acid  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: protein  
US-08-615-170-3

Query Match 6.3%; Score 70.5; DB 1; Length 432;  
Best Local Similarity 23.3%; Pred. No. 8;  
Matches 38; Conservative 20; Mismatches 62; Indels 43; Gaps 7;  
  
Qy 3 VLSAPGLGRFRIL-GLRSSYGPVAVQA-----RGVHQSVA--TDGPSSTQPALPKAR 50  
Db 137 IISATAFHSKMAPLGLPRSAYPVSGFWOGALPGQAGSSQDVKPTQOQYALQPSLPLG 196  
Qy 51 AVAP-----KPSRGEYVVAKLDLVNWARSSLWMTFGLACCAVEMMHMAAPRYD 102  
Db 197 FDSPTGLPPSSSTPAWQGRVAS-----SKLWMLEFS----AFLEQQQDDQDTYN 241  
Qy 103 MDRFGVVFASPRQSDVMIVAGLTINKMAPALRKVYDQMPER 145  
Db 242 KHLFVHIGQSNPSYSDPYL-----EAVDIRQIYDKPEKK 276

RESULT 7  
US-08-615-170-5  
; Sequence 5, Application US/08615170  
; Patent No. 5776776  
; GENERAL INFORMATION:  
; APPLICANT: ORDAHL, Charles P.  
; APPLICANT: AZAKIE, Anthony  
; APPLICANT: MAR, Janet H.  
; APPLICANT: FARRANCE, Iain K.G.  
; APPLICANT: HALL, Deborah E.  
; APPLICANT: STEWART, Alexandre F.R.  
; APPLICANT: LARKIN, Sarah B.  
; TITLE OF INVENTION: DTEP-1 ISOFORMS AND USES THEREOF  
; NUMBER OF SEQUENCES: 32

;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: Townsend and Townsend Kourie and Crew  
;; STREET: Stewart Street Tower, One Market Plaza  
;; CITY: San Francisco  
;; STATE: California  
;; COUNTRY: US  
;; ZIP: 94105-1493  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: PatentIn Release #1.0, Version #1.25  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/615,170  
;; FILING DATE:  
;; CLASSIFICATION: 435  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: PCT/US95/01526  
;; FILING DATE: 06-FEB-1995  
;; CLASSIFICATION: 435  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 08/191,493  
;; FILING DATE: 04-FEB-1994  
;; CLASSIFICATION: 435  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Heslin, James M.  
;; REGISTRATION NUMBER: 29,541  
;; REFERENCE/DOCKET NUMBER: 2307U-053120  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (415) 326-2400  
;; TELEFAX: (415) 326-2422  
;; INFORMATION FOR SEQ ID NO: 5:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 445 amino acids  
;; TYPE: amino acid  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: protein  
US-08-615-170-5

Query Match 6.3%; Score 70.5; DB 1; Length 445;  
Best Local Similarity 23.3%; Pred. No. 8.3;  
Matches 38; Conservative 20; Mismatches 62; Indels 43; Gaps 7;  
  
Qy 3 VLSAPGLGRFRIL-GLRSSYGPVAVQA-----RGVHQSVA--TDGPSSTQPALPKAR 50  
Db 150 IISATAFHSKMAPLGLPRSAYPVSGFWOGALPGQAGSSQDVKPTQOQYALQPSLPLG 209  
Qy 51 AVAP-----KPSRGEYVVAKLDLVNWARSSLWMTFGLACCAVEMMHMAAPRYD 102  
Db 210 FDSPTGLPPSSSTPAWQGRVAS-----SKLWMLEFS----AFLEQQQDDQDTYN 254  
Qy 103 MDRFGVVFASPRQSDVMIVAGLTINKMAPALRKVYDQMPER 145  
Db 255 KHLFVHIGQSNPSYSDPYL-----EAVDIRQIYDKPEKK 289

RESULT 8  
US-08-258-287B-37  
; Sequence 37, Application US/08258287B  
; Patent No. 6083735  
; GENERAL INFORMATION:  
; APPLICANT: Yuan, Junying  
; APPLICANT: Miura, Masayuki  
; TITLE OF INVENTION: Programmed Cell Death Genes and Proteins  
; NUMBER OF SEQUENCES: 85  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox  
; STREET: 1100 New York Avenue, Suite 600  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: USA  
; ZIP: 20005



```

; TELEFAX: 508-872-5415
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4302 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-658-136-5

Query Match      6.38; Score 70; DB 3; Length 4302;
Best Local Similarity 23.38; Pred. No. 2.8e+02;
Matches 58; Conservative 24; Mismatches 57; Indels 110; Gaps 16;

QY 3 VLSAPGLGRILGRSSVFAVGARVGHOSVATDGPSTQPALPKARAVAPKPSRGEY 62
DB 707 VMLPG-----DLVGLQHDAGPGAL---LHCSFAPGHPG-----PRAPVLSANASSWLPH 753
QY 63 VYAKDDLNVNARRSSLWPTFGLACCAVEMHMAAP-----RYDM 103
DB 754 LPAQLE-----GTW-----ACPACALRLIAATEQLTVLLGLRPNPGLRLPGRYEV 798
QY 104 D-----RGVVFRAASPROSDVMIVAGILTNKMAPALRKVYDQMPER- 145
DB 799 RAEVNGVSRNLSGSDVV---SP-----LVYI---PAPRDGRL 836
QY 146 YVSMGSC-----ANG-----GGYHYYSVVRGCDRVP--VDIYIPGCPPTAE 188
DB 837 YVPTNGSALVQVDSGANATATARWPG-----SVSARENVCPALVATFVPGCPWEIN 890
QY 189 ALLYGILOL 197
DB 891 DTLFSVAL 899

RESULT 11
US-08-443-260-3
; Sequence 3, Application US/08443260
; Patent No. 5942234
; GENERAL INFORMATION:
; APPLICANT: RALSTON, ROBERT O.
; APPLICANT: MARCUS, FRANK
; APPLICANT: THUDUM, KENT B.
; APPLICANT: GERVASE, BARBARA A.
; APPLICANT: HALL, JOHN A.
; TITLE OF INVENTION: HEPATITIS C VIRUS ANTIGENIC PROTEINS
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Chiron Corporation
; STREET: 4560 Horton Street
; CITY: Emeryville
; STATE: California
; COUNTRY: USA
; ZIP: 94608-2916
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/443,260
; FILING DATE: 17-MAY-1995
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: HARBIN, ALISA A.
; REGISTRATION NUMBER: 33,895
; REFERENCE/DOCKET NUMBER: 0154.006
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (510) 601-2708
; TELEFAX: (510) 655-3542
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2955 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; NAME/KEY: Modified-site
; LOCATION: 1949
; OTHER INFORMATION: /note= "There is a heterogeneity at this location; Xaa = Glu or Gly"
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 1948
; OTHER INFORMATION: /note= "There is a heterogeneity at this location; Xaa = His or Leu"
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 1877
; OTHER INFORMATION: /note= "There is a heterogeneity at this location; Xaa = Cys or Tyr"
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 1471
; OTHER INFORMATION: /note= "There is a heterogeneity at this location; Xaa = Ser or Thr"
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 1276
; OTHER INFORMATION: /note= "There is a heterogeneity at this location; Xaa = Leu or Pro"
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 1117
; OTHER INFORMATION: /note= "There is a heterogeneity at this location; Xaa = Ser or Thr"
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 1114
; OTHER INFORMATION: /note= "There is a heterogeneity at this location; Xaa = Pro or Ser"
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 848
; OTHER INFORMATION: /note= "There is a heterogeneity at this location; Xaa = Asn or Tyr"
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 603
; OTHER INFORMATION: /note= "There is a heterogeneity at this location; Xaa = Ile or Leu"
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 334
; OTHER INFORMATION: /note= "There is a heterogeneity at this location; Xaa = Met or Val"
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 176
; OTHER INFORMATION: /note= "There is a heterogeneity at this location; Xaa = Ile or Thr"
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 11
; OTHER INFORMATION: /note= "There is a heterogeneity at this location; Xaa = Arg or Lys"
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 9
; OTHER INFORMATION: /note= "There is a heterogeneity at this location; Xaa = Arg or Lys"
; FEATURE:
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein

```



```

; OTHER INFORMATION: this location; Xaa = Leu or Pro"
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 1454
; OTHER INFORMATION: /note= "There is a heterogeneity at
; OTHER INFORMATION: this location; Xaa = Cys or Tyr"
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 1471
; OTHER INFORMATION: /note= "There is a heterogeneity at
; OTHER INFORMATION: this location; Xaa = Ser or Thr"
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 1877
; OTHER INFORMATION: /note= "There is a heterogeneity at
; OTHER INFORMATION: this location; Xaa = Glu or Gly"
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 1948
; OTHER INFORMATION: /note= "There is a heterogeneity at
; OTHER INFORMATION: this location; Xaa = His or Leu"
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 1949
; OTHER INFORMATION: /note= "There is a heterogeneity at
; OTHER INFORMATION: this location; Xaa = Cys or Ser"
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 2021
; OTHER INFORMATION: /note= "There is a heterogeneity at
; OTHER INFORMATION: this location; Xaa = Gly or Val"
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 2349
; OTHER INFORMATION: /note= "There is a heterogeneity at
; OTHER INFORMATION: this location; Xaa = Ser or Thr"
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 2385
; OTHER INFORMATION: /note= "There is a heterogeneity at
; OTHER INFORMATION: this location; Xaa = Phe or Tyr"
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 2386
; OTHER INFORMATION: /note= "There is a heterogeneity at
; OTHER INFORMATION: this location; Xaa = Ala or Ser"
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 2502
; OTHER INFORMATION: /note= "There is a heterogeneity at
; OTHER INFORMATION: this location; Xaa = Phe or Leu"
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 2690
; OTHER INFORMATION: /note= "There is a heterogeneity at
; OTHER INFORMATION: this location; Xaa = Gly or Arg"
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 2921
; OTHER INFORMATION: /note= "There is a heterogeneity at
; OTHER INFORMATION: this location; Xaa = Arg or Gly"
;
; US-08-442-805A-3

```

```

Query Match      6.2%; Score 69.5; DB 3; Length 2955;
Best Local Similarity 20.8%; Pred. No. 1.8e+02;
Matches 48; Conservative 29; Mismatches 75; Indels 79; Gaps 14;

QY 7 PGLRGRFLLG---LRSSVGPVQARGVHVSATDGPSSSTQPALPKARAVAPKPSRGEYV 63
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 25 PG--GGQIVGGVLLPRRGLRGLVTRKTSRSQPRGRQPIPKAR----RPGR----- 74
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 64 VAKLDDLNVNARRSSLWPMTFGLACCAVEMHMAAPRYDMRFGVVFASPRQ----- 116

```

```

Db 75 -----TWAQPGYPMPL-YGNEGCGWAGW-LLSPRGRSRPSWG-----PTDPRRRNRNLGK 121
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 117 -----SDVM-----IVAGTLTNKMAPALRKVD-----QMPE---PRYVYSM 150
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 122 VIDLTGCFADLMGYIPLVGAPLGSARALAHGVRLVDGVVYATGNLPGCSFSXFLAL 181
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 151 GSC-----ANGGGYHY-----SYSVVRGCDRIVPVDIYIPGCPP 185
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 182 LSCLTVPASAYQVRNSTGLYHTNDPCNPSSIVYEADAI-----LHTPGCVP 228
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 13
US-08-443-900A-3
; Sequence 3, Application US/08443900A
; Patent No. 6074852
; GENERAL INFORMATION:
; APPLICANT: RALSTON, ROBERT O.
; APPLICANT: MARCUS, FRANK
; APPLICANT: THUDUM, KENT B.
; APPLICANT: GERVAISE, BARBARA A.
; APPLICANT: HALL, JOHN A.
; TITLE OF INVENTION: HEPATITIS C VIRUS ASIALOGLYCOPROTEINS
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Chiron Corporation
; STREET: 4560 Horton Street
; CITY: Emeryville
; STATE: California
; COUNTRY: USA
; ZIP: 94608-2916
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; FILING DATE: 17-May-1995
; CLASSIFICATION: 435
; APPLICATION NUMBER: US/08/443.900A
; ATTORNEY/AGENT INFORMATION:
; NAME: HARBIN, ALISA A.
; REGISTRATION NUMBER: 33,895
; REFERENCE/DOCKET NUMBER: 0154.004
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (510) 601-2708
; TELEFAX: (510) 655-3542
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2955 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 9
; OTHER INFORMATION: /note= "There is a heterogeneity at
; OTHER INFORMATION: this location; Xaa = Arg or Lys"
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 11
; OTHER INFORMATION: /note= "There is a heterogeneity at
; OTHER INFORMATION: this location; Xaa = Asn or Thr"
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 176
; OTHER INFORMATION: /note= "There is a heterogeneity at
; OTHER INFORMATION: this location; Xaa = Ile or Thr"
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 334
; OTHER INFORMATION: /note= "There is a heterogeneity at

```



RESULT 15  
 US-08-249-843-3  
 ; Sequence 3, Application US/08249843  
 ; Patent No. 6274148  
 ; GENERAL INFORMATION:  
 ; APPLICANT: RALSTON, ROBERT O.  
 ; APPLICANT: MARCUS, FRANK  
 ; APPLICANT: THUDIOM, KENT B.  
 ; APPLICANT: GERVASE, BARBARA A.  
 ; APPLICANT: HALL, JOHN A.  
 ; TITLE OF INVENTION: HEPATITIS C VIRUS ASIALOGLYCOPROTEINS  
 ; NUMBER OF SEQUENCES: 3  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Chiron Corporation  
 ; STREET: 4560 Horton Street  
 ; CITY: Emeryville  
 ; STATE: California  
 ; COUNTRY: USA  
 ; ZIP: 94608-2916  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk

```

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/249,843
FILING DATE: 26-MAY-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: HARDIN, ALISA A.
REGISTRATION NUMBER: 33,895
REFERENCE/DOCKET NUMBER: 0154,003
TELECOMMUNICATION INFORMATION:
TELEPHONE: (510) 601-2708
TELEFAX: (510) 655-3542
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 2955 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
FEATURE:
NAME/KEY: Modified-site
LOCATION: 9
OTHER INFORMATION: /note= "There is a heterogeneity at
OTHER INFORMATION: this location; Xaa = Arg or Lys"
FEATURE:
NAME/KEY: Modified-site
LOCATION: 11
OTHER INFORMATION: /note= "There is a heterogeneity at
OTHER INFORMATION: this location; Xaa = Asn or Thr"
FEATURE:
NAME/KEY: Modified-site
LOCATION: 176
OTHER INFORMATION: /note= "There is a heterogeneity at
OTHER INFORMATION: this location; Xaa = Ile or Thr"
FEATURE:
NAME/KEY: Modified-site
LOCATION: 334
OTHER INFORMATION: /note= "There is a heterogeneity at
OTHER INFORMATION: this location; Xaa = Met or Val"
FEATURE:
NAME/KEY: Modified-site
LOCATION: 603
OTHER INFORMATION: /note= "There is a heterogeneity at
OTHER INFORMATION: this location; Xaa = Ile or Leu"
FEATURE:
NAME/KEY: Modified-site
LOCATION: 848
OTHER INFORMATION: /note= "There is a heterogeneity at
OTHER INFORMATION: this location; Xaa = Asn or Tyr"
FEATURE:
NAME/KEY: Modified-site
LOCATION: 1114
OTHER INFORMATION: /note= "There is a heterogeneity at
OTHER INFORMATION: this location; Xaa = Pro or Ser"
FEATURE:
NAME/KEY: Modified-site
LOCATION: 1117
OTHER INFORMATION: /note= "There is a heterogeneity at
OTHER INFORMATION: this location; Xaa = Ser or Thr"
FEATURE:
NAME/KEY: Modified-site
LOCATION: 1276
OTHER INFORMATION: /note= "There is a heterogeneity at
OTHER INFORMATION: this location; Xaa = Leu or Pro"
FEATURE:
NAME/KEY: Modified-site
LOCATION: 1454
OTHER INFORMATION: /note= "There is a heterogeneity at
OTHER INFORMATION: this location; Xaa = Cys or Tyr"
FEATURE:
NAME/KEY: Modified-site

```

```

? LOCATION: 1471
? OTHER INFORMATION: /note= "There is a heterogeneity at
? OTHER INFORMATION: this location; Xaa = Ser or Thr"
? FEATURE:
? NAME/KEY: Modified-site
? LOCATION: 1877
? OTHER INFORMATION: /note= "There is a heterogeneity at
? OTHER INFORMATION: this location; Xaa = Glu or Gly"
? FEATURE:
? NAME/KEY: Modified-site
? LOCATION: 1948
? OTHER INFORMATION: /note= "There is a heterogeneity at
? OTHER INFORMATION: this location; Xaa = His or Leu"
? FEATURE:
? NAME/KEY: Modified-site
? LOCATION: 1949
? OTHER INFORMATION: /note= "There is a heterogeneity at
? OTHER INFORMATION: this location; Xaa = Cys or Ser"
? FEATURE:
? NAME/KEY: Modified-site
? LOCATION: 2021
? OTHER INFORMATION: /note= "There is a heterogeneity at
? OTHER INFORMATION: this location; Xaa = Gly or Val"
? FEATURE:
? NAME/KEY: Modified-site
? LOCATION: 2349
? OTHER INFORMATION: /note= "There is a heterogeneity at
? OTHER INFORMATION: this location; Xaa = Ser or Thr"
? FEATURE:
? NAME/KEY: Modified-site
? LOCATION: 2385
? OTHER INFORMATION: /note= "There is a heterogeneity at
? OTHER INFORMATION: this location; Xaa =Phe or Tyr"
? FEATURE:
? NAME/KEY: Modified-site
? LOCATION: 2386
? OTHER INFORMATION: /note= "There is a heterogeneity at
? OTHER INFORMATION: this location; Xaa = Ala or Ser"
? FEATURE:
? NAME/KEY: Modified-site
? LOCATION: 2502
? OTHER INFORMATION: /note= "There is a heterogeneity at
? OTHER INFORMATION: this location; Xaa = Phe or Leu"
? FEATURE:
? NAME/KEY: Modified-site
? LOCATION: 2690
? OTHER INFORMATION: /note= "There is a heterogeneity at
? OTHER INFORMATION: this location; Xaa = Gly or Arg"
? FEATURE:
? NAME/KEY: Modified-site
? LOCATION: 2921
? OTHER INFORMATION: /note= "There is a heterogeneity at
? OTHER INFORMATION: this location; Xaa = Arg or Gly"
? US-08-249-843-3
?
? Query Match 6.2%; Score 69.5; DB 4; Length 2955
? Best Local Similarity 20.8%; Pred. No. 1.8e+02;
? Matches 48; Conservative 29; Mismatches 75; Indels 79
?
? QY 7 PGLRGFRILG--LRSSVGPAVOARGVHQSVADGDSSQTCPALPKARAVAPKPSFS
? | | .|.|. | | | | | | | | | | | | | | | | | | | |
? 25 PG--GGQIVGGVILLPRGRLGLVTRATKTSERQSRGRRQPIPKAR----RPEGH
?
? QY 64 VAKLDDLNNWARRSSLLPMTFGLACCACAVEMHMAAPRYDMDFRGVFRASPRQ--
? | | | | | | | | | | | | | | | | | | | | | | | |
? Db 75 -----TWAQPGYWPML-YGNEGCGWAGW-LLSPRGSRPSWG---PTDPRRRSS
?
? QY 117 -----SDVM-----IVAGTITNKMAPALRKVYD-----QMPE---PRY
? | | | | | | | | | | | | | | | | | | | | | | | |
? Db 122 VIDTLTCGADFLMGYIPLVGAPJGSAARALAHGVRLDGVNVATGNLPGCSFSFX
?
? QY 151 GSC-----ANGGGVYHY-----SYSVVRGCGDRIVPVDIVIPGCPP 185

```



Db 182 USCLTVPASAYQVRNSTGLYHVTNDCPNSSIVEADAI-----LHTPGCVP 228

Search completed: February 12, 2002, 20:40:47  
Job time: 2862 sec



alignment\_scores:  
 Quality: 1118.00 Length: 213  
 Ratio: 5.249 Gaps: 0  
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:  
 US-09-525-867-1 x AAI60566 ..

Align seg 1/1 to: AAI60566 from: 1 to: 752

1 MetAlaValLeuSerAlaProGlyLeuArgGlyPheArgIleLeuGlyLe 17  
 |||  
 5 ATGGCGGTGCTGCTGAGTCTCTGCGCTCGCGGCTTCGGATCCTTGGTCT 54  
 |||  
 17 uArgSerSerValGlyProAlaValGlnAlaArgGlyValHisGlnSerV 34  
 |||  
 55 CGCGTCCAGCGTGGCGCGGCTGTGCAGGCACGAGGTGTCATCAGAGCG 104  
 |||  
 34 aAlaThrAspGlyProSerSerThrGlnProAlaLeuProLysAlaArg 50  
 |||  
 105 TGGCCACGATGGCCCAAGCAGCACCAGCTGCCCTGCCAAGGCCGAGA 154  
 |||  
 51 AlaValAlaProLysProSerSerArgGlyGlyValValAlaLysLe 67  
 |||  
 155 GCCGTGGCTCCCAACCCAGCAGCGCGGGGAGTATGTGTGSCCAAGCT 204  
 |||  
 67 uAspAspLeuValAsnTrpAlaArgArgSerSerLeuTrpProMetThrP 84  
 |||  
 205 GGATGACCTCTCAACTGGGCGCGCGGAGTTCTCTGTGGCCCATGACCT 254  
 |||  
 84 heGlyLeuAlaCysCysAlaValGluMetMetHisMetAlaAlaProArg 100  
 |||  
 255 TCGGCTGGCTGTGCGCGGTGGAGATGACATGGCAGCACCCTCCG 304  
 |||  
 101 TyrAspMetAspArgPheGlyValValPheArgAlaSerProArgGlnse 117  
 |||  
 305 TACGACATGGACCGCTTTGGCGTGGTCTTCGCGCGCAGCGCGCGCAGTC 354  
 |||  
 117 rAspValMetIleValAlaGlyThrLeuThrAsnLysMetAlaProAlaL 134  
 |||  
 355 CGAGCTCATGATCGTGGCGCGGCACACTCACCAACAAGATGGCCCGCAG 404  
 |||  
 134 euArgLysValTyrAspGlnMetProGluProArgTyrValValSerMet 150  
 |||  
 405 TTCGCAAGGTCTACGACACAGATGCGGAGCGCGGTACGTGTCTCCATG 454  
 |||  
 151 GlySerCysAlaAsnGlyGlyTyrTyrTyrHisTyrSerTyrSerValVa 167  
 |||  
 455 GGGAGCTGGCCCAACGAGGAGGAGGCTACTACCATATTCCTACTCGTGT 504  
 |||  
 167 lArgGlyCysAspArgIleValProValAspIleTyrIleProGlyCysP 184  
 |||  
 505 GAGGGGCTGCGACCGCATCTGTCGCGGTGGACATCTACATCCCGAGGTGCC 554  
 |||  
 184 roProThrAlaGluAlaLeuLeuTyrGlyIleLeuGlnLeuArgLys 200  
 |||  
 555 CACCTACGGCGGAGGCGCTCTCTACGGCATCTCGAGCTGCGAGGGAAG 604  
 |||  
 201 lIleLysArgGluArgArgLeuGlnIleTrpTyrArgArg 213  
 |||  
 605 ATCAAGCGGGAGCGGAGGCTGCAGATCTGTGTACCGCAGG 643

seq\_name: /SDS2/gcgdata/geneseq/geneseq/NA2001.DAT:AAI58780

seq\_documentation\_block:

ID AAI58780 standard; cDNA; 772 BP.

XX AC AAI58780;

XX DT 22-OCT-2001 (first entry)

XX

DE Human polynucleotide SEQ ID NO 983.  
 XX Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;  
 KW peripheral nervous system; neuropathy; central nervous system; CNS;  
 KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;  
 KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;  
 KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;  
 KW leukaemia; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200153312-A1.  
 XX  
 PD 26-JUL-2001.  
 XX  
 PF 26-DEC-2000; 2000WO-US34263.  
 XX  
 PR 21-JAN-2000; 2000US-0488725.  
 PR 25-APR-2000; 2000US-0552317.  
 PR 09-JUL-2000; 2000US-0598042.  
 PR 19-JUL-2000; 2000US-0620312.  
 PR 03-AUG-2000; 2000US-0653450.  
 PR 14-SEP-2000; 2000US-0662191.  
 PR 19-OCT-2000; 2000US-0693036.  
 PR 29-NOV-2000; 2000US-0727344.  
 XX  
 PA (HYSE-) HYSEQ INC.  
 XX  
 PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;  
 PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;  
 PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;  
 XX  
 DR WPI: 2001-442253/47.  
 DR P-PSDB; AAM39524.  
 XX  
 PT Novel nucleic acids and polypeptides, useful for treating disorders  
 PT such as central nervous system injuries -  
 XX  
 PS Claim 1; SEQ ID NO 983; 10078pp; English.  
 XX  
 CC The invention relates to human nucleic acids (AAI57798-AAI61369) and  
 CC the encoded polypeptides (AAM38642-AAM42213) with nootropic,  
 CC immunosuppressant and cytostatic activity. The polynucleotides are useful  
 CC in gene therapy. A composition containing a polypeptide or polynucleotide  
 CC of the invention may be used to treat diseases of the peripheral nervous  
 CC system, such as peripheral nervous injuries, peripheral neuropathy and  
 CC localised neuropathies and central nervous system diseases, such as  
 CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic  
 CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the  
 CC utilisation of the activities such as: Immune system suppression,  
 CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic  
 CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,  
 CC assays for receptor activity, arthritis and inflammation, leukaemias and  
 CC C.N.S disorders.  
 CC Note: The sequence data for this patent did not form part of the printed  
 CC specification.  
 XX  
 SQ Sequence 772 BP; 136 A; 265 C; 245 G; 126 T; 0 other;

alignment\_scores:  
 Quality: 1118.00 Length: 213  
 Ratio: 5.249 Gaps: 0  
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

US-09-525-867-1 x AAI58780 ..

Align seg 1/1 to: AAI58780 from: 1 to: 772

1 MetAlaValLeuSerAlaProGlyLeuArgGlyPheArgIleLeuGlyLe 17  
 |||  
 24 ATGGCGGTGCTGCTGAGTCTCTGCGCTCGCGGCTTCGGATCCTTGGTCT 73



```
268 GCTGGCGCGTGGAGATGACATGCAGCAGCCCGCTACGACATGGAC 317
105 ArgPheGlyValValPheArgAlaSerProArgGlnSerAspValMetIi 121
318 CGCTTTGGCGTGGTCTTCGCGCGCAGCCCGCCAGTCGACGTCATGAT 367
121 evalAlaGlyThrLeuThrAsnLysMetAlaProAlaLeuArgLysValT 138
368 CQTGGCCCGGCACACTCCACCAAGAATGGCCCGCCAGCGCTTCGAAGTCT 417
138 yAspGlnMetProGluProArgTyrValValSerMetGlySerCysAla 154
418 AGCACCATGCGCGGAGCGCGCTACGTGCTCTCCATGGGAGACTGGCC 467
155 AsnGlyGlyTyrTyrHisTyrSerTyrSerValValArgLysCysAs 171
468 AACGGAGAGGCTACTACCACTATCTACTCTCGTGGTGGAGGCTGCCGA 517
171 PArgIleValProValAspIleTyrIleProGlyCysProProThrAlaG 188
518 CGCATCTGTCGCGTGGACATCTACATCCAGGCTGCCACTACGGCCG 566
188 luAlaLeuLysTyrGlyLeuGlnLeuGlnArgLysIleLysArgGlu 204
567 AGGCCCTGCTCTACGGCATCTCGAGCTGCAGAGGAGGAATCAAGCGGGAG 616
205 ArgArgLeuGlnIleTyrTyrArgArg 213
617 CGGAGGCTGCAGATCTGTGATCCGCGAGG 643

seq_name: /SIDS2/gcgdata/geneseq/geneseqn/NA2000.DAT:AAc43639
seq_documentation_block:
ID_AAc43639 standard; DNA; 911 BP.
XX
AC AAC43639;
XX
DT 17-OCT-2000 (first entry)
XX
DE Zea mays DNA fragment SEQ ID NO: 39953.
XX
KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway; metabolic;
KW pathway; promoter; termination sequence; corn; ss.
XX
OS Zea mays subsp. mays.
XX
XX EP1033405-A2.
XX
PD 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-0301439.
XX
XX 25-FEB-1999; 99US-0121825.
XX 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132407.
PR 05-MAY-1999; 99US-0132484.
PR 06-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
106-MAY-1999; 99US-0132487.
07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
```

```
PR 22-JUL-1999; 99US-01451192.
PR 23-JUL-1999; 99US-01451145.
PR 23-JUL-1999; 99US-01452118.
PR 23-JUL-1999; 99US-01452224.
PR 26-JUL-1999; 99US-01452276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-01459118.
PR 27-JUL-1999; 99US-01459119.
PR 28-JUL-1999; 99US-0145911.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147433.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149358.
PR 17-AUG-1999; 99US-01493175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 30-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.

PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161932.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

alignment_scores:
    Quality: 737.50      Length: 214
    Ratio: 4.120        Gaps: 3
    Percent Similarity: 83.645      Percent Identity: 66.822

alignment_block:
US-09-525-867-1 x AAC43639 ..

Align seg 1/1 to: AAC43639 from: 1 to: 911

7 ProGlyLeuArgGlyPheArgIleLeuGlyLeuArgSerSerValGlyPr 23
   ||| ||||| : : : ||||| : : : ||||| : : : |||||
65 C C T C G T C C C A C G A G G T G A G A A C C C T A G C A C C G A C C A T G G C G T G C T C C 114
   ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : |||
23 O A L a V a l G l n A l a A r G l y . V a l . . . . . H i s G l n S e r 33
   ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : |||
115 C G C G C A C C G C G G T G C C C T C C T T C C A C C C G C G C G T A C T C C G C T 164
   ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : |||
34 V a l A l a T h r A s p G l y P r o S e r S e r T h r G l n P r o A l a L e u P r o L y s A l a A r 50
   ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : |||
165 G C G G T G C T G C G G G C C C T C C C C A G C G T C C C A G C G . . . C C A T A C G G G G 211
   ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : |||
50 G A L a V a l A l a P r o L y s P r o S e r S e r A r g G l y L u t y r V a l V a l A l a L y s L 67
   ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : |||
212 C G C C C C C C A C C G G C G A T G C G A A G A C G C G A G T C G T G T C T C T A A G G 261
   ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : |||
67 e u A s p A s p L e u V a l A s n T r p A l a A r g A r g S e r S e r L e u T r p P r o M e t h r 83
   ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : |||
262 T G G A C G A C C T G A T G A C T G G G C G T A A G G G C T G A T C T G C G C C A T G A C C 311
   ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : |||
84 P h e G l y L e u A l a C y s A l a V a l G l u M e t H i s M e t A l a A l a P r o A r 100
   ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : |||
312 T T C G G C T C G C T G T C G C G G T G G A T G A T G A C A C C G C G C G C G T C C C G 361
   ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : |||
100 g T y r A s p M e t A s p A r g P h e G l y V a l V a l P h e A r g A l a S e r P r o A r g G l n S 117
   ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : |||
362 C T A C G A C T T C G A C C G G T T C G G G T C A T C T C C G T C C G T C G C G C G C A G T 411
   ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : |||
117 e r A s p V a l M e t I l e V a l A l a G l y T h r L e u T h r A s n L y s M e t A l a P r o A l a 133
   ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : |||
412 C C G A T T G C A T G A T C G T C C G C G C A C G C T C A C C A C A G A T G G C T C C A G C C 461
   ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : |||
134 L e u A r g L y s V a l T y r A s p G l n M e t P r o G l u P r o A r g T y r V a l V a l S e r M e 150
   ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : |||
462 C T C C C A A G G T T A T G A C C A A T G C T G A G C C T A G T G G G T A T T T C A A T 511
   ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : |||
150 t G l y S e r C y s A l a A s n G l y G l y T y r T y r H i s T y r S e r T y r S e r V a l V 167
   ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : |||
512 G G C A G C T G T G C C A A C G G T G G T G G A T A C T A C C A T T A C T C C T A C T C T G T T G 561
   ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : |||
167 a l A r g G l y C y s A s p A r g I l e V a l P r o V a l A s p I l e T y r I l e P r o G l y C y s 183
   ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : |||
562 T A C G T G G A T G T A C C G T A T A G T C C T C T G T G G A C A T C A C T C C C T G G G T G C 611
   ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : |||
184 P r o P r o T h r A l a G l u A l a L e u L e u T y r G l y I l e L e u G l n L e u G l n A r g L y 200
   ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : |||
612 C C A C C A A C T G C T G A G G C A C T G C T G T A C G C G C T T C T C A G C T C C A A A A G A A 661
   ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : |||
```

```
200 sileLysArgGluArgLeuGlnIleTrpTyrArgArg 213
|||||:|||||:|||||:|||||:|||||:|||||:
662 GATCAACAGGGGTAAAGGATTTCCTTCACTGTTGGACCAAG 701

seq_name: /SIBS2/gcgdata/geneseq/geneseqn/NA2000.DAT.AAC36686
seq_documentation_block:
ID AAC36686 standard; DNA; 937 BP.
XX AC AAC36686;
XX DT 17-OCT-2000 (first entry);
XX DE Arabidopsis thaliana DNA fragment SEQ ID NO: 14712.
XX KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway;
KW metabolic pathway; promoter; termination sequence; ss.
XX OS Arabidopsis thaliana.
XX PN EPI033405-A2.
XX PD 06-SEP-2000.
XX PF 25-FEB-2000; 2000EP-0301439.
XX PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 21-JUN-1999; 99US-0139763.
PR 22-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 24-JUN-1999; 99US-0140354.
PR 28-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 28-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145276.
PR 26-JUL-1999; 99US-0145224.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 09-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
```

16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 14-OCT-1999; 99US-0159638.  
PR 18-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160768.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 21-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160980.  
PR 22-OCT-1999; 99US-0160981.  
PR 22-OCT-1999; 99US-0160983.  
PR 25-OCT-1999; 99US-0161404.  
PR 25-OCT-1999; 99US-0161405.  
PR 25-OCT-1999; 99US-0161406.  
PR 26-OCT-1999; 99US-0161359.  
PR 26-OCT-1999; 99US-0161360.  
PR 26-OCT-1999; 99US-0161361.  
PR 28-OCT-1999; 99US-0161920.  
PR 28-OCT-1999; 99US-0161992.  
PR 28-OCT-1999; 99US-0161993.  
PR 29-OCT-1999; 99US-0162144.

alignment\_scores: Quality: 725.00 Length: 198  
Ratio: 4.290 Gaps: 2  
Percent Similarity: 85.354 Percent Identity: 68.182

alignment\_block:

US-09-525-867-1 x AAC36686 ..

Align seg 1/1 to: AAC36686 from: 1 to: 937

24 AlaValGlnAlaArgGlyVal.....HisGlnSerValAlaThrAs 37

110 GCGCTCGCGCTGCGCTGCTCTCACATCCACATCCCTTCCGCGCTCT 159  
37 pGlyProSerSer.....ThrGlnProAlaLeuProLysA 49  
160 CTCCTCCCTCCATACCCAGCTTCTACACAGACCTGGTCTCTCTCGA 209  
49 laArgAlaValAlaProLysProSerSerArgGlyGlyValValAla 65  
210 CCGCCCT 259  
66 LysLeuAspLeuValAsnTrpAlaArgArgSerSerLeuTrpProme 82  
260 AGGTGATGATCTCATGAACCTGGCTGCTACCGGATGATCTGGCCAT 309  
82 tThrPheGlyLeuAlaCysCysAlaValGluMetMethHisMetAlaAla 99  
310 GACCTTTGGCTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 359  
99 roArgTyrAspMetAspArgPheGlyValValPheArgAlaSerPro 115  
360 CTGCTTACGATCTCGATCGATTCGGTATCATCTTCAGGCTAGCTCGC 409  
116 GlnSerAspValMetIleValAlaGlyThrLeuThrAsnLysMetAla 132  
410 CAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 459  
132 oAlaLeuArgLysValTyrAspGlnMetProGluProArgTyrValVal 149  
460 GGCTCTCGCAAGTTTATGACAGATGCTGAGCAAGTGGTGATTT 509  
149 erMetGlySerCysAlaAsnGlyGlyGlyTyrTyrHisTyrSerTyr 165  
510 CAAATGGAGTTGTGCAATGGAGTGGATGATGATGATGATGATGATG 559  
166 ValValArgGlyCysAspArgIleValProValAspIleTyrIlePro 182  
560 GTGTTTCGAGGATGTGACAGAAATGTCCAGTCGACATATAGTCCCG 609  
182 yCysProProThrAlaGluAlaLeuTyrGlyIleLeuGlnLeuGlnA 199  
610 GTGCCCAACCAACCGCTGAGCTTTGCTCTATGACTACTCCAGCTTC 659  
199 rGlySileLysArgGluArgArgLeuGlnIleTyrTyrArgArg 213  
660 AGAAATCAACAGCGCAAGGATTTCTTCTGATTGGTGGAAACAAG 703

seq\_name: /SIDS2/gcgdata/geneseq/geneseq/NA2000.DAT:AAC38037

seq\_documentation\_block:

ID AAC38037 standard; DNA; 976 BP.

XX AAC38037;

AC AAC38037;

DT 17-OCT-2000 (first entry)

XX Arabidopsis thaliana DNA fragment SEQ ID NO: 19562.

XX Hybridisation assay; genetic mapping; gene expression control;  
XX protein identification; signal transduction pathway;  
XX metabolic pathway; promoter; termination sequence; ss.

OS Arabidopsis thaliana.

XX EP1033405-A2.

XX PD 06-SEP-2000.

XX PF 25-FEB-2000; 2000EP-0301439.

XX PR 25-FEB-1999; 99US-0121825.

PR 05-MAR-1999; 99US-0123180.

PR 09-MAR-1999; 99US-0123548.



PR 23-MAR-1999; 99US-0125788.  
PR 25-MAR-1999; 99US-0126264.  
PR 29-MAR-1999; 99US-0126785.  
PR 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128234.  
PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0128845.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 28-APR-1999; 99US-0130891.  
PR 28-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 30-APR-1999; 99US-0132407.  
PR 04-MAY-1999; 99US-0132484.  
PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 07-MAY-1999; 99US-0132487.  
PR 11-MAY-1999; 99US-0132863.  
PR 14-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135253.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 25-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0137222.  
PR 01-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 14-JUN-1999; 99US-0138847.  
PR 16-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 17-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 21-JUN-1999; 99US-0139763.  
PR 22-JUN-1999; 99US-0139817.  
PR 23-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140354.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140698.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141844.  
PR 01-JUL-1999; 99US-0142055.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143544.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 09-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147935.  
PR 09-AUG-1999; 99US-0147935.  
PR 11-AUG-1999; 99US-0148171.  
PR 12-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.



PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 21-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.

PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 11-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159325.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 14-OCT-1999; 99US-0159638.  
PR 18-OCT-1999; 99US-0159584.  
PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160768.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 21-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160980.  
PR 22-OCT-1999; 99US-0160981.  
PR 22-OCT-1999; 99US-0160989.  
PR 25-OCT-1999; 99US-0161404.  
PR 25-OCT-1999; 99US-0161405.  
PR 25-OCT-1999; 99US-0161406.  
PR 26-OCT-1999; 99US-0161359.  
PR 26-OCT-1999; 99US-0161360.  
PR 26-OCT-1999; 99US-0161361.  
PR 28-OCT-1999; 99US-0161920.  
PR 28-OCT-1999; 99US-0161992.  
PR 28-OCT-1999; 99US-0161993.

```

seq_name: /SIDS2/gcgdata/geneseq/geneseqn/NA2000.DAT:AAA81468

seq_documentation_block:
ID   AAA81468 standard; DNA; 72750 BP.
XX
XX
AC   AAA81468;
XX
XX
04-DEC-2000 (first entry)
XX
XX
DE   N. meningitidis partial DNA sequence gnm_16 SEQ ID NO:16.
XX
XX
KW   Neisseria meningitidis; Neisseria gonorrhoeae; genome; immunogenic;
KW   antigen; vaccine; diagnosis; infection; antibacterial; identification;
53826 GTTGTGGCCGGTTACTTTCGGTTGCGCTGCGCGGTGGAATGATGC 53777

```





PS Claim 1; Page 263-277; 545pp; English.

XX The present invention relates to a Moraxella catarrhalis genomic library comprising of a combination of 41 nucleic acid molecules (see AAF28514-AAF28554). The library has a number of uses described in the specification e.g. is useful for identifying diagnostic and therapeutic compositions. M. catarrhalis (Branhamella catarrhalis) is a large aerobic, gram-negative diplococcus, normally found among the bacterial flora of human upper airways. M. catarrhalis is known to cause acute, localised infections such as otitis media, sinusitis and bronchopulmonary infection and life-threatening, systemic diseases including endocarditis and meningitis.

XX Sequence 58909 BP; 16632 A; 11699 C; 13199 G; 17379 T; 0 other;

alignment\_scores:  
Quality: 431.50 Length: 177  
Ratio: 3.104 Gaps: 5  
Percent Similarity: 78.531 Percent Identity: 49.718

alignment\_block:  
US-09-525-867-1 x AAF28543 ..

Align seg 1/1 to: AAF28543 from: 1 to: 58909

36 ThrAspGlyProSerThrGlnProAlaLeuProLysAlaArgAlaVal 52  
26810 ACCAACATACCAAGCAATACCCACAAAGCCGCGAAATCATTCAGCAT 26859  
52 lalaProLysProSer..SerArgGlyGluTyrValValAlaLysLeuAs 68  
26860 ATCACCAAGCCGAGGTGGATAAATGTTTT...TTGGGCAGATTGTC 26906  
68 pAsp.....LeuValAsnTrpAlaArgSerSerLeuTrp 81  
26907 CGATTTCAGCACAACCTTCGCCAAGTGGGTGCGAAAGAACTCTTATGCG 26956  
81 roMetThrPheGlyLeuAlaCysCysAlaValGluMetMethHisMetAla 97  
26957 CCTTAACTTGGCACAAGCTGTGTATGTGGAGATGCCACGACCTTG 27006  
98 AlaProArgTyrAspMetAspArgPheGlyVal...ValPheArgAlaSe 113  
27007 ACAGCGCTGATGATTGTCTGCTTTTGGGGCAGAGAGTCAATTCGTGCTC 27056  
113 rProArgLysSerAspValMetIleValAlaGlyThrLeuThrasnLysM 130  
27057 CCTTCGCCAGCTGATGATGATGATGATGCGGGGACTGTTTGTGCAAAA 27106  
130 etAlaProAlaLeuArgLysValTyrAspGlnMetProGluProArgTyr 146  
27107 TGGCGCTGTGATTCAGGCTGTATGAACAAATGTAGAACCTAAATGG 27156  
147 ValValSerMetGlySerCysAlaAsnGlyGlyTyrTyrHisTyrSe 163  
27157 GTCACTCCATGGGGCGGTGCGCAACTCTGTGGCATGACGACATT.. 27204  
163 rTyrSerValValArgGlyCysAspArgIleValProValAspIleTyrI 180  
27205 .TATTCGGTGGTGCAGGGGTGGATAAATCTTCCTGTGGATGTGATG 27253  
180 leProGlyCysProProThrAlaGluAlaLeuLeuTyrGlyIleLeuGln 196  
27254 TTCGGGGGTGTCGCCCGACCTGAGGCACTCAATCAAGCCATCATGCTA 27303  
197 LeuGlnArgLysIleLysArgGluArgArg 206  
27304 TTACAAGAGAGCATTCAAAAGAGACGCCGT 27333

seq\_name: /SIDS2/9c9data/geneseq/geneseq/NA1990.DAT.AAQ04525  
seq\_documentation\_block:

ID AAQ04525 standard; DNA; 134525 BP.

XX AAQ04525;

XX 01-OCT-1990 (first entry)

XX Total base sequence of rice plant chloroplast DNA.

XX Chloroplast; true grass; rice plant; ss.

XX Oryza sativa.

XX JP02100682-A.

XX 12-APR-1990.

XX 07-OCT-1988; 88JP-0251967.

XX 07-OCT-1988; 88JP-0251967.

XX (MITK) MITSUI TOATSU CHEM INC.

XX WPI; 1990-159709/21.

XX Chloroplast DNA of true grasses - used to produce various DNA base sequences by decomposition of rice plant DNA.

XX Claim 1; Fig 1; 20pp; Japanese.

XX The sequence is that of the whole of rice chloroplast DNA.

XX Sequence 134525 BP; 41249 A; 26129 C; 26331 G; 40816 T; 0 other;

alignment\_scores:  
Quality: 427.50 Length: 148  
Ratio: 3.562 Gaps: 1  
Percent Similarity: 81.081 Percent Identity: 49.324

alignment\_block:  
US-09-525-867-1 x AAQ04525/rev ..

Align seg 1/1 to reverse of: AAQ04525 from: 1 to: 134525

58 SerArgGlyGluTyrValValAlaLysLeuAspLeuValAsnTrpAl 74  
49210 AGTTCATTCGGTATTTCACATACACTAAAGAGCTTTCGAATTGGTC 49161  
74 aArgArgSerSerLeuTrpProMetThrPheGlyLeuAlaCysCysAla 91  
49160 AAGACTCTCTAGTTATGGCCCTCTATATGTCACAGTGTGTGTTTCA 49111  
91 aGluMetMethHisMetAlaProArgTyrAspMetAspArgPheGly 107  
49110 TTGAATTTGCTTCATTAATAGGCTCAGATTCGATTCGTTAATGA 49061  
108 ValValPheArgAlaSerProArgGlnSerAspValMetIleValAla 124  
49060 TTGGTACCAAGATCAAGTCCCTAGCAAGCGGACCTAATTTTAAACGCGG 49011  
124 yThrLeuThrasnLysMetAlaProAlaLeuArgLysValTyrAspGln 141  
49010 TACAGTAACAATGAAATGGCTCCCTCTTTAGTGAGATTATAGCAAA 48961  
141 etProGluProArgTyrValValSerMetGlySerCysAlaAsnGly 157  
48960 TGCCTGAACCAAAATACGTTCATTCCTATGGAGCTGTACTATTACTGT 48911  
158 GlyTyrTyrHisTyr...SerTyrSerValValArgGlyCysAspArg 173  
48910 GGAATGTTTCAGTACGAGTTCCTATAGTACTGTTCCGGGAGTGTAGAGT 48861  
173 eValProValAspIleTyrIleProGlyCysProProThrAlaGluAla 190

48860 AATTCCTGTGGAGCTTACTTCCGGTGTCCACCTAAACAGAGGCTG 48811  
190 euleuTyrrGlylleLeuGlnLeuGlnArgLysileLysArgGlu 204  
48810 TTATAGATGCCCTAACAAACTCGTGAAGATATCGGAGAA 48767

seq\_name: /SIDS2/gcgcdata/geneseq/geneseq/NA2000.DAT: AAC95299

seq\_documentation\_block:  
ID AAC95299 standard; cDNA; 446 BP.

AC AAC95299;

DT 19-FEB-2001 (first entry)

DE Cat flea head and nerve cord (HNC) cDNA, SEQ ID NO:1794.

KW Cat flea; head and nerve cord nucleic acid; HNC; flea infestation;  
KW vaccine; antiparasitic; therapeutic target; diagnosis; detection; ss.  
OS Ctenocephalides felis.

XX WO200061621-A2.

XX 19-OCT-2000.

XX 07-APR-2000; 2000WO-US09437.

XX 09-APR-1999; 99US-0128704.

XX (HESK-) HESKA CORP.

XX Brandt KS, Gaines PJ, Stinchcomb DT, Wisniewski N;

XX WPI; 2000-656323/63.

XX Flea Malpighian tubule and head and nerve cord tissue derived nucleic  
PT acids useful for the prevention, diagnosis and treatment of flea  
PT infestations -

XX Claim 26; Page 848; 964pp; English.

XX The invention relates to novel cat flea (Ctenocephalides felis) nucleic  
CC acids which are expressed in hindgut and Malpighian tubule (HMT) tissue  
CC or head and nerve cord (HNC) tissue. The invention also relates to the  
CC encoded proteins. The invention additionally encompasses expression  
CC constructs, recombinant viruses and recombinant cells comprising the  
CC nucleic acids of the invention, recombinant production of the proteins,  
CC antibodies against the proteins, a method of identifying inhibitors of  
CC the proteins, and compositions comprising the inhibitors for  
CC administration to an animal. The nucleic acids, and the proteins they  
CC encode may be used in the prevention, treatment and diagnosis of diseases  
CC associated with flea infestations. For example, the nucleic acids may be  
CC used to produce an HMT or HNC protein according to standard recombinant  
CC DNA methodology by inserting the nucleic acids into a host cell and  
CC culturing the cell to express the protein. The HMT and HNC nucleic acids  
CC may also be used as DNA probes in diagnostic assays (e.g., PCR) to detect  
CC and quantitate the presence of cat flea or other homologous nucleic acid  
CC sequences in samples. They may also be used to study the expression and  
CC function of the proteins and their role in metabolism. The HMT and HNC  
CC proteins may be used as antigens in the production of specific  
CC antibodies, and in assays to identify modulators (agonists and  
CC antagonists) of HMT and/or HNC protein expression and activity. The  
CC anti-HMT/HNC protein antibodies and antagonists may also be used to  
CC downregulate protein expression and activity. The antibodies may also be  
CC used as diagnostic agents for detecting the presence of flea polypeptides  
CC in samples (e.g., by enzyme linked immunosorbent assay (ELISA)). The  
CC present sequence represents a cat flea HNC cDNA of the invention.

XX Sequence 446 BP; 179 A; 52 C; 75 G; 139 T; 1 other;

alignment\_scores:  
Quality: 418.00 Length: 84  
Ratio: 5.160 Gaps: 0  
Percent Similarity: 96.429 Percent Identity: 86.905  
alignment\_block:  
US-09-525-867-1 x AAC95299 ..  
Align seg 1/1 to: AAC95299 from: 1 to: 446  
130 MetAlaProAlaLeuArgLysValTyrrAspGlnMetProGluProArgTy 146  
2 ATGGCTCCAGCGCTCGAAAAGTTTATGATCAATGCCCTGACACCGTTG 51  
146 rValValSerMetGlySerCysAlaAsnGlyGlyTyrrHisTyrrs 163  
52 GGTATCTCTATGGTAGTTGTGCTAATGGAGGAGGTATTACCATATT 101  
163 erTyrrSerValValArgGlyCysAspArgGlyLeuValProValAspIleTyrr 179  
102 CATATCTCTAGTAGGGGTTGTGATAGAAATAGTACCAGTTGACATATAT 151  
180 IleProGlyCysProProThrAlaGluAlaLeuLeuTyrrGlyIleLeu 196  
152 GTACCAGGATGCTCCCAACTGCAGAAAGCTTTGTTATATGTTGTTTACA 201  
196 nLeuGlnArgLysLysArgGluArgArgLeuGlnIleTyrrTyrrArgA 213  
202 GCTTCAAAAGAAAGTTAAACGAATGAAACCTGCAATGCGGTATAGAA 251  
213 rg 213  
252 AA 253

seq\_name: /SIDS2/gcgcdata/geneseq/geneseq/NA1998.DAT: AAV61275

seq\_documentation\_block:

ID AAV61275 standard; cDNA; 789 BP.

AC AAV61275;

DT 06-JAN-1999 (first entry)

DE 3' cDNA sequence of prostate tumour clone J1-21.

XX Prostate; cancer; tumour; vaccine; immunogen; clone; ss.

OS Homo sapiens.

XX WO9837093-A2.

XX 27-AUG-1998.

XX 25-FEB-1998; 98WO-US03492.

XX 09-FEB-1998; 98US-0020956.

XX 25-FEB-1997; 97US-0806099.

XX 01-AUG-1997; 97US-0904804.

XX (CORI-) CORIAX CORP.

XX Dillon DC, Xu J;

XX WPI; 1998-609886/51.

XX Polypeptides comprising immunogenic portions of prostate proteins -  
PT used in a vaccine for the treatment of prostate cancer

XX Claim 12; Page 48; 130pp; English.

XX The present sequence is a DNA which encodes an immunogenic portion  
CC of a prostate tumour protein. The encoded immunogen, or the DNA itself,  
CC can be used as a vaccine for the treatment of prostate cancer. The DNA





```

|| .....|||
464 TTGGGCGCGCCCGCCAGTCTCAGTATGATGTGCGNGGNANNTA 415
127 ThrAsnLysMetAlaProAlaLeuArgLysValTyrAspGlnMetProG1 143
||:||||| |||||:||||| ||||| ||||| |||||
414 ACCAAAAGAGAGGCCCGCCAGNGTTGC.AAGGTTTAAGACCAGATGCCGGA 366
143 uProArgTyrValValSerMetGlySerCysAlaAsnGlyGlyTyrT 160
||| ||||| ||||| ||||| ||||| ||||| |||||
365 GCCGCGANAAGTGGTTCCANGGGGAGTGN.GCCAACGGNGGGGTAAN 317
160 yrHisTyrSerTyrSerValValArgGlyCysAspArgIleValProVal 176
||| ||||| :||| ||||| ||||| ||||| ||||| |||||
316 ACCACTATTCTTAATNGTGTGTGGGCGNTGGACCCGCAATTGTGCCCGTG 267
177 AspIleTyrIleProGlyCysProProThrAlaGluAlaLeuLeuTyrG1 193
:||||| ||||| ||||| |||||:||||| ||||| ||||| |||||
266 GAAATNAAATCCAGGTCGCCACCNANGCCGAGGCCCTGTTTANGG 217
193 yIleLeuGlnLeuGlnArgLysIleLysArgGluArgArgLeuGlnIleT 210
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
216 CATCTGCAGCTGCAGAGGAANATCAACGGGAGCGGAGGCTGCANATT 167
210 rpTyrArgArg 213
|||||
166 GGTACCGCAGG 156
```

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 12, 2002, 16:51:49 ; Search time 1347.47 Seconds  
(without alignments)  
10088.296 Million cell updates/sec

Title: US-09-525-867-9  
Perfect score: 824  
Sequence: 1 cgcctcagcgcgtcgcgcg.....cggcgcaaaaaaaaaaaaaa 824

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 1472140 seqs, 824859755 residues

Total number of hits satisfying chosen parameters: 2944280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.\*

- 1: gb\_ba.\*
- 2: gb\_hgt.\*
- 3: gb\_in.\*
- 4: gb\_om.\*
- 5: gb\_ov.\*
- 6: gb\_pat.\*
- 7: gb\_ph.\*
- 8: gb\_pl.\*
- 9: gb\_pr.\*
- 10: gb\_ro.\*
- 11: gb\_sts.\*
- 12: gb\_sy.\*
- 13: gb\_un.\*
- 14: gb\_vi.\*
- 15: em\_ba.\*
- 16: em\_fun.\*
- 17: em\_hum.\*
- 18: em\_in.\*
- 19: em\_om.\*
- 20: em\_or.\*
- 21: em\_ov.\*
- 22: em\_pat.\*
- 23: em\_ph.\*
- 24: em\_pl.\*
- 25: em\_ro.\*
- 26: em\_sts.\*
- 27: em\_sy.\*
- 28: em\_un.\*
- 29: em\_vi.\*
- 30: em\_htgo\_hum.\*
- 31: em\_htgo\_inv.\*
- 32: em\_htgo\_rod.\*
- 33: em\_htg\_hum.\*
- 34: em\_htg\_inv.\*
- 35: em\_htg\_rod.\*
- 36: em\_htg\_other.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

\*

Result No.	Score	Query Match	Length	DB ID	Description
1	754.8	91.6	787	9	BC005954 Homo sapi
2	738.2	89.6	794	9	BC001715 Homo sapi
3	704	85.4	754	9	AF060512 Homo sapi
4	458.8	60.5	752	4	BTPTSSTU
c 5	384.6	46.7	789	6	AX106251 Sequence
c 6	384.6	46.7	789	6	AX140542 Sequence
c 7	308	37.4	83607	2	AC019889
8	308	37.4	175124	3	AC007856
9	308	37.4	224100	3	AE003770
10	308	37.4	311043	3	AC009888
11	280.4	34.0	3739	1	PDEMOUVR
12	279.2	33.9	29563	2	AC014169
13	279.2	33.9	140685	2	AC010919
14	279.2	33.9	163378	2	AC008336
c 15	279.2	33.9	327446	3	AE003500
c 16	276	33.5	329709	1	AP002997
17	274.2	33.3	2274	8	YLI250340
18	273.8	33.2	6229	1	RME245398
19	273.8	33.2	299350	1	SME591786
20	269.2	32.7	19614	1	AF029365
21	268.4	32.6	911	8	BONADHPST
22	268.4	32.6	951	8	NCIRONSUL
c 23	261.6	31.7	10337	1	AE005868
24	255.4	31.0	897	8	AF281035
25	251.8	30.6	951	8	ATNADRUO
26	250.6	30.4	935	8	STPSST
27	250.6	30.4	1021	8	AF185271
28	248	30.1	1620	1	AF055637
29	240	29.1	13592	1	AE008054
30	233	28.3	15312	1	AE003884
c 31	210.2	25.5	340806	1	NMA1Z2491
32	209	25.4	4787	3	PARMTICFL
33	208.6	25.3	10467	1	AE002381
34	208.6	25.3	349980	6	AX043922
35	199.4	24.2	14886	1	TAU52917
36	199.4	24.2	40469	3	MPAPEN
c 37	197	23.9	312430	1	RPXX02
c 38	194.4	23.6	59034	3	AF007261
39	192.8	23.4	48063	8	AF288090
40	190	23.1	34875	9	AC005329
41	188	22.8	39739	1	SCD16A
42	187.6	22.8	20293	1	SCD20
c 43	186.4	22.6	45223	8	AF110138
c 44	177.4	21.5	14695	1	AE007270
45	177.4	21.5	18681	1	RME245399

ALIGNMENTS

RESULT	1				
BC005954	BC005954	787 bp	mRNA	PRI	12-JUL-2001
LOCUS	Homo sapiens				
DEFINITION	Homo sapiens, clone MGC:14592				IMAGE:4276489, mRNA, complete cds.
ACCESSION	BC005954				
VERSION	BC005954.1				GI:13543602
KEYWORDS	MGC.				
SOURCE	human.				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;				
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
TITLE	1 (bases 1 to 787)				
JOURNAL	Strausberg,R.				
	Direct Submission				
	Submitted (02-APR-2001) National Institutes of Health, Mammalian				
	Gene Collection (MGC), Cancer Genomics Office, National Cancer				
	Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,				
	USA				
REMARK	NIH-MGC Project URL: http://mgc.nci.nih.gov				
COMMENT	Contact: MGC help desk				



```
LOCUS AF060512 754 bp mRNA PRI 02-JAN-2001
DEFINITION Homo sapiens clone 016d03 My017 protein mRNA, complete cds.
ACCESSION AF060512
VERSION AF060512.1 GI:12001973
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 754)
AUTHORS Mao,X.M., Xie,Y., Huang,X.Y., Ying,K. and Dai,J.L.
TITLE Direct Submission
JOURNAL Submitted (20-APR-1998) Institute of Genetics, School of Life
Science, Fudan University, 220 Handan Rd., Shanghai 200433,
P.R.China
FEATURES
source 1..754
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="016d03"
/tissue_type="brain"
/dev_stage="fetus"
10..654
CDS
/codon_start=1
/product="My017 protein"
/protein_id="AAG43130.1"
/db_xref="GI:12001974"
/translation="MAVLSAPGLRGLRILGRSSVGLAVQARGVHQSVATDGPSSTOP
ALPKRAVAPKSSRGEEYVAKLDLVLNWRSSLWPMFTGLACCAVEMHMAAPRYD
MDRFYVFRASPRQSDVMIVAGTLTKMAPALRKSPRPDAGALRHLGELRQGRLL
LFILGLGRLRPHRARGHLRPLTTEALLYILOLQKIKRRRLQIWR"
BASE COUNT 120 a 267 c 239 g 128 t
ORIGIN
Query Match 85.4%; Score 704; DB 9; Length 754;
Best Local Similarity 98.7%; Pred. No. 3.2e-104;
Matches 732; Conservative 0; Mismatches 5; Indels 5; Gaps 2;
Qy 70 gagccagaatgagcggtgtgtcagctctcgctgcgctgcggtctccggtctctgtgtctg 129
Db 1 GAGCCAAAGATGCGGTGTGTCTAGCTCTCGCTCGCTCGCGGTCTCGGATCTTGGTCTG 60
Qy 130 cgtctccagctggtgcccgcgtgtgcaggcacgaggtgtccatcagagcgtggtccaccgat 189
Db 61 CGCTCCACGCTGGGCTGTGTGTCAGGCACGAGGTGTCTCATCAGAGCGTGCCACCGAT 120
Qy 190 ggcacaagcagcagcagcctgcctgccaaaggccagagcagcgtggtcccaaacccagc 249
Db 121 GGGCCAAAGCAGCCAGCTGCCCTGCCAAAGCCAGAGCGGTGCTCCCAACCCAGC 180
Qy 250 agccggggcagtagtggtggccaaagctggatgaacctgcaactggccgcccggagt 309
Db 181 AGCGGGGCGAGTATGTGTGGCGCAAGCTGGATGACCTCGTCAACTGGCGCGCGGAGT 240
Qy 310 tctctggtgcccagatgaacctcggcctgctgctggtgagcgtgagatgatgcacatgca 369
Db 241 TCTGTGTGGCCCATAGACCTTCGGCTGTGTGCTGTGTGCGGTGGAGATGATGACATGGCA 300
Qy 370 gcaacccgcgtacgacatggaccgcttggcgtgtcttcgcgcagcagccgcgcagtc 429
Db 301 GCACCCCGCTACGACATGACCGCTTTGGCGTGTCTTCCGCGCCAGCGCCGCGAGTCC 360
Qy 430 gacgtctatgctgtgcccggcacactcaccacaagaatggccccagcgttgcgaaggt- 488
Db 361 GACGTCTATGATCTGTGCGCGGCACACTCACCAACAGATGCCCGCAGCTTCGCAAGTCT 420
Qy 489 ---ctacgaccagatgccggagcgcgcgtacgtggtctccatggggagcgtgcgcaacgg 545
Db 421 CCCCCAGCAGCATGCCGGAGCGCGCTACGTGGTCTCCATGGGGAGCTGGCCCAACGG 480
Qy 546 agagggtactaccactattctactcgtgtgtgtaggggtgcgaccgcatcgtgtccgt 605
AF060512
```





```

RESULT 7
AC019889/c
LOCUS AC019889 83607 bp DNA HTG 03-JAN-2000
DEFINITION Drosophila melanogaster, *** SEQUENCING IN PROGRESS ***, in ordered
pieces.
ACCESSION AC019889
VERSION AC019889.1 GI:6665008
KEYWORDS HTG; HTGS_PHASE2.
SOURCE fruit fly.
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 83607)
AUTHORS Adams,M. and Venter,J.C.
TITLE Direct Submission
JOURNAL Submitted (30-DEC-1999) Celera Genomics, 45 West Gude Drive,
Rockville, MD, USA
COMMENT This sequence was identified as CDM:10211288 by the submitter.
For more information on this record e-mail to flycelera.com.
* NOTE: This is a 'working draft' sequence.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
FEATURES
Source
1..83607
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
BASE COUNT 23093 a 18268 c 19032 g 23214 t
ORIGIN
Query Match 37.4%; Score 308; DB 2; Length 83607;
Best Local Similarity 59.8%; Pred. No. 2.6e-41;
Matches 416; Conservative 0; Mismatches 180; Indels 0; Gaps 0;
Qy 127 ctgcgtccagctggcccggtgtgcaggccagcaggtgtccatcagagctggccacc 186
|||||
Db 14363 CTGCTCGCAGTGAAGTCTGCAATCCCAAGTTCGTGGCGCAGAACCAATGCGGTG 14304
Qy 187 gatggcccaagcagcaccagctgcctgcctgcacaaagccagacgctggtcccaacc 246
|||||
Db 14303 TACGATTACCGTGGCCAAATCCGCAAGTGGGGTACTCGGCCTTCGGACGCAACAG 14244
Qy 247 agcagcggggcagatgtgtgtggccaaagctggtgacatcgtcaactggccgcgcgg 306
|||||
Db 14243 AAGACCTGGGCGAGTGGACCTGTGCCAGGTGGATGACCTGTAACCTGGGACGCAAG 14184
Qy 307 agttctctgtggccatgaccttgcgctgtgctgctgctgctgctgctgctgctgctg 366
|||||
Db 14183 GGATCCCTCTGGCGCGTACCTTGTGCTGCGCTGTGCGCGCTGAAATGATGACATT 14124
Qy 367 gcagcaccgctacgacatggaccgctttggcgtggtcttccggcgccagccgcgcgag 426
|||||
Db 14123 CGGCTCTCTGCTACGACATGATGCTACGCGTGTGCTGCTGCGACATCTCTCGCCAG 14064
Qy 427 tccagctcatatgtctgtggccgacactccacaaagatggcccgcttgcgaag 486
|||||
Db 14063 CGGACATGCTCATCTGTGGCGCGCAACCTGACCAACAGATGCACCGGCTTCGGGAG 14004
Qy 487 gtctacgacagatgcgagcgagcgtcagctgtgtctcatggagctgcgcgaacagga 546
|||||
Db 14003 ATCTACACACAGATGCCGACGCGAGATGGGTATTTCCATGGGCGATTCGCCAATGTT 13944
Qy 547 ggaggtactaccactattctactcgtgtgtgtgaggggtgtgcagcgcgctgtgccctgt 606
|||||
Db 13943 GCGGCTCTACTACCACTACTCTACTCGTGTGCGGGTGTGGGATCGCATCTTTCGGGTG 13884
Qy 607 gacatctacatccagctgccacatcagccgagcgcctgtctctacgacatcctgag 666
|||||
Db 13883 GACATCTACGTGCGCGGATGTGCGCCACCGCGAGGCTTAAATGTACGGAATCTGTGAG 13824
Qy 667 ctgcagaggaagatcaagcggagcggaggtgtgcagatctgtgtacgcgaggtagcg 722

```

```

Db 13823 CTGCAAGAAGCTAAGCGCATGAGGACCTGCTGCTAGAGGAGTAGAG 13768
|||||
RESULT 8
AC007856 175124 bp DNA INV 17-FEB-2001
LOCUS AC007856
DEFINITION Drosophila melanogaster, chromosome 3R, region 99B-99B, BAC clone
BAC05L03, complete sequence.
ACCESSION AC007856
VERSION AC007856.7 GI:12957650
KEYWORDS HTG.
SOURCE fruit fly.
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 175124)
AUTHORS Celiker,S.E., Adams,M.D., Kronmiller,B., Tyler,D., Wan,K.H.,
Holt,R.A., Evans,C.A., Gocayne,J.D., Amanatides,P.G., Brandon,R.C.,
Rogers,Y., An,H., Baldwin,B., Banzon,J., Beeson,K.Y., Busam,D.A.,
Carlson,J.W., Center,A., Champe,M., Davenport,L.B., Dietz,S.M.,
Dodson,K., Dorsett,V., Doup,L.E., Doyle,C., Dresnek,D., Farfan,D.,
Ferriera,S., Frise,E., Galle,R.F., Garg,N.S., George,R.A.,
Gonzalez,M., Houck,J., Hoskins,R.A., Hostin,D., Howland,T.J.,
Ibegwam,C., Jalali,M., Kruse,D., Li,P., Mattei,B., Moshrefi,A.,
McIntosh,T.C., Moy,M., Murphy,B., Nelson,C., Nelson,K.A., Nunoo,J.,
Paclet,J., Paragas,V., Park,S., Patel,S., Pfeiffer,B., Scheeler,F.,
Phouanavong,S., Pittman,G.S., Puri,V., Richards,S., Scheeler,F.,
Stapleton,M., Strong,R., Svirskas,R., Tector,C., Williams,S.M.,
Zaveri,J.S., Smith,H.O., Rubin,G.M. and Venter,J.C.
TITLE Sequencing of Drosophila chromosome 3R, region 99B-99B
Unpublished
2 (bases 1 to 175124)
Celiker,S.E., Agbayani,A., Arcaina,T.T., Baxter,E., Blazej,R.G.,
Butenhoff,C., Champe,M., Chavez,C., Chew,M., Ciesiolka,L.,
Doyle,C.M., Farfan,D.E., Galle,R., George,R.A., Harris,N.L.,
Hoskins,R.A., Houston,K.A., Hummasti,S.R., Karra,K., Kearney,L.,
Kim,E., Lee,B., Lewis,S., Li,P., Lomotan,M.A., Mazda,P.,
Moshrefi,A.R., Moshrefi,M., Nixon,K., Paclet,J.M., Park,S.,
Pfeiffer,B., Poon,L., Sequeira,A., Sethi,H., Snir,E.,
Svirskas,R.R., Wan,K.H., Weinburg,T., Zhang,R., Zieran,L.L. and
Rubin,G.M.
Direct Submission
Submitted (17-JUN-1999) Drosophila Genome Center, Lawrence Berkeley
Laboratory, MS 64-121, Berkeley, CA 94720, USA
On Feb 17, 2001 this sequence version replaced gi:6633905.
Sequence submitted by:
Berkeley Drosophila Genome Project
Lawrence Berkeley National Laboratory, MS 64-121
Berkeley, CA 94720
This sequence was assembled using end sequences from a whole genome
shotgun and from subclones of this BAC and its neighboring clones.
For further information about this sequence, including its location
and relationship to other sequences, please visit our sequence
archive Web site (http://www.fruitfly.org/sequence/) or send email
to bdgp@fruitfly.berkeley.edu.
FEATURES
Location/Qualifiers
1..175124
/organism="Drosophila melanogaster"
/strain="y; cn bw sp"
/db_xref="taxon:7227"
/chromosome="3R"
/map="99B-99B"
/clone="BAC05L03 (D795)"
/clone_lib="RPCI-98 (Roswell Park Cancer Institute
Drosophila melanogaster BAC library, partial EcoRI in
PBAC3.6)"
BASE COUNT 47809 a 39206 c 39325 g 48784 t
ORIGIN
Query Match 37.4%; Score 308; DB 3; Length 175124;

```



Best Local Similarity 69.8%; Pred. No. 1.9e-41;  
Matches 416; Conservative 0; Mismatches 180; Indels 0; Gaps 0;

QY 127 ctgcgctcagcggcggccgcgtgctcagcagcaggtgtccatcagcagcgtggtgccaacc 186  
|||||  
DB 22093 CTGCGTCGACGTGGAGTCTGCATCCAGTGGTCTGTCGCCAGAGACCATGCGGGTG 22152  
|||||  
QY 187 gatgcccacaagcagcagccagccctgcctgcacaaagccagagccgtggtcccaaaccc 246  
|||||  
DB 22153 TACGATTACCGCTGCCAAATCCGCCAAAGTGGGCTACTCGGCTTCGGACGCAACCCAG 22212  
|||||  
QY 247 agcaccggggcagtagtgtgtgagcaagctgatgataacatctcaactggtgcccaccgg 306  
|||||  
DB 22213 AAGACTGGGCGAGTGCACCTGTGCAGGTGTGATGATGCTTAACCTGGGACCAAG 22272  
|||||  
QY 307 agttctgtgtggccatgacctctggcctggtcgtgctgcgcctgtggagatgatgcacatg 366  
|||||  
DB 22273 GGATCCCTCTGCGCGCTGACCTTTGGTCTGCGCTGCGCGCTGCGAAATGATGCACATT 22332  
|||||  
QY 367 gcagcaccgccctacgacatgagccgtttggcgtgtgtcttcgacccagccgcgcag 426  
|||||  
DB 22333 GCGGCTCTCGCTACGACATGATGCTACGCTGTGTGTTCGAGCATCTCTCGCCAG 22392  
|||||  
QY 427 tcgcagctcatgatcgtggcggcgacactcaccacaaagatggcccgagcgttcgcaag 486  
|||||  
DB 22393 GCGGACGTGCTCATCTGCGCGGAACCTGACCAACAAGATGACACGGGCTTTGGAAG 22452  
|||||  
QY 487 gtctacgacagatgcggagccgcgtacgtgtgtctccatggggagctgcgcacaagga 546  
|||||  
DB 22453 AFTCTACGACAGATGCCGAGCCGAGATGGGTCTATTTCCATGGGCGATTCGCGCAATGTT 22512  
|||||  
QY 547 gtagctactaccacttctactcgtgtgtgtgaggggtgcgacgcagcgtggtccgctg 606  
|||||  
DB 22513 GCGGCTACTACCACTACTCTACTCTGCTGCTGCGGCTGCGATCGCATGTTCGGGTG 22572  
|||||  
QY 607 gacatctacatccagcgtgcccacccagcagcagcagcagcagcagcagcagcagcag 666  
|||||  
DB 22573 GACATCTACGTGCGCGGATGTCGCCACCGCCGAGGCTTAATGTACGGAATCTCTGCAG 22632  
|||||  
QY 667 ctgcagagagatcaagcgggagcagcagcagcagcagcagcagcagcagcagcagcag 722  
|||||  
DB 22633 CTGCAGAGAGAGTAAGCCGATGAGGACCTCGAGATGTGGTACAGGAAGTAGAG 22688  
|||||

RESULT 9  
AE003770  
LOCUS AE003770 224100 bp DNA INV 05-OCT-2000  
DEFINITION Drosophila melanogaster genomic scaffold 14200001386035 section 95  
of 105, complete sequence.  
ACCESSION AE003770 AE002708  
VERSION AE003770.1 GI:7301790  
KEYWORDS HTG.  
SOURCE fruit fly.  
ORGANISM Drosophila melanogaster  
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;  
Muscophora; Ephydroidea; Drosophilidae; Drosophila.  
1 (bases 1 to 224100)  
ADAMS M.D., Celniker S.E., Holt R.A., Evans C.A., Gccayne, J.D.,  
Amanatides, P.G., Scherer S.E., Li, P.W., Hoskins R.A., Galle, R.F.,  
George, R.A., Lewis S.E., Richards, S., Ashburner, M., Henderson, S.N.,  
Sutton, G.G., Wortman, J.R., Yandell, M.D., Zhang, Q., Chen, L.X.,  
Brandon, R.C., Rogers, Y.H., Blazej, R.G., Champe, M., Pfeiffer, B.D.,  
Wan, K.H., Doyle, C., Baxter, E.G., Helt, G., Helt, G., Nelson, C.R., Gabor  
Miklos, G.L., Abril, J.F., Agbayani, A., An, H.J.,  
Andrews-Pfannkoch, C., Baldwin, D., Ballew, R.M., Basu, A.,  
Baxendale, J., Bayraktaroglu, L., Beasley, E.M., Beeson, K.Y.,  
Benos, P.V., Berman, B.P., Bhandari, D., Bolshakov, S., Borkov, D.,  
Botchan, M.R., Bouck, J., Brokstein, P., Brottier, P., Burtis, K.C.,  
Busam, D.A., Butler, H., Cadieu, E., Center, A., Chandra, I.,  
Cherry, J.M., Cawley, S., Dahlke, C., Davenport, L.B., Davies, P., de  
Pablos, B., Delcher, A., Deng, Z., Mays, A.D., Dew, I., Dietz, S.M.,  
Dodson, K., Dou, L.E., Downes, M., Dugan-Rocha, S., Dunkov, B.C.,

Dunn, P., Durbin, K.J., Evangelista, C.C., Ferraz, C., Ferreira, S.,  
Fleischmann, W., Fowler, C., Gabriellian, A.E., Garg, N.S.,  
Gibart, W.M., Glasser, K., Glodek, A., Gong, F., Gorrell, J.H., Gu, Z.,  
Guan, P., Harris, M., Harris, N.L., Harvey, D., Heiman, T.J.,  
Hernandez, J.R., Houck, J., Hostin, D., Houston, K.A., Howland, T.J.,  
Wei, M.H., Ibegwam, C., Jaitani, M., Kalush, F., Karpen, G.H., Ke, Z.,  
Kenshin, J.A., Ketchum, K.A., Kinmel, B.E., Kodira, C.D., Kraft, C.,  
Kravitz, S., Kulp, D., Lai, Z., Lasko, P., Lei, Y., Levitsky, A.A.,  
Li, J., Li, Z., Liang, Y., Lin, X., Liu, X., Mattel, B., McIntosh, T.C.,  
McLeod, M.P., McPherson, D., Merkulov, G., Milshina, N.V., Mobarry, C.,  
Morris, J., Moshrefi, A., Mount, S.M., Moy, M., Murphy, B., Murphy, L.,  
Muzny, D.M., Nelson, D.L., Nelson, D.R., Nelson, K.A., Nixon, K.,  
Nusskern, D.R., Paclet, J.M., Palazzi, M., Picman, G.S., Pan, S.,  
Pollard, J.V., Puri, V., Reese, M.G., Reinert, K., Remington, K.,  
Saunders, R.D., Scheeler, F., Shen, H., Shue, B.C., Siden-Kianos, I.,  
Simpson, M., Skupski, M.P., Smith, T., Spier, E., Spradling, A.C.,  
Stapleton, M., Strong, R., Sun, E., Svirskas, R., Tector, C., Turner, R.,  
Venter, E., Wang, A.H., Wang, X., Wang, Z.Y., Wasserman, D.A.,  
Weinstock, G.M., Weissbach, J., Williams, S.M., Woodage, T.,  
Worley, K.C., Wu, D., Yang, S., Yao, Q.A., Ye, J., Yeh, R.F.,  
Zaveri, J.S., Zhan, W., Zhang, G., Zhao, Q., Zheng, L., Zheng, X.H.,  
Zhong, F.N., Zhong, W., Zhou, X., Zhu, S., Zhu, X., Smith, H.O.,  
Gibbs, R.A., Myers, E.W., Rubin, G.M. and Venter, J.C.  
The genome sequence of *Drosophila melanogaster*  
Science 287 (5461), 2185-2195 (2000)  
20196006  
2 (bases 1 to 224100)  
Adams, M.D., Celniker, S.E., Gibbs, R.A., Rubin, G.M. and Venter, J.C.  
Direct Submission  
Submitted (21-04-2000) Celera Genomics, 45 West Gude Drive,  
Rockville, MD, USA  
FEATURES  
1. 224100  
Location/Qualifiers  
/organism="Drosophila melanogaster"  
/db\_xref="taxon:7227"  
/chromosome="3R"  
/gene="CG2014"  
/product="CT6471"  
/db\_xref="FLYBASE:FBan0002014"  
/db\_xref="FLYBASE:FBgn0039669"  
/gene="CG2014"  
/map="99B3-99B4"  
/db\_xref="FLYBASE:FBan0002014"  
/db\_xref="FLYBASE:FBgn0039669"  
6113..6751  
/gene="CG2014"  
/note="CG2014" gene product"  
/codon\_start=1  
/db\_xref="FLYBASE:FBan0002014"  
/db\_xref="FLYBASE:FBgn0039669"  
/protein\_id="AAF56901.1"  
/db\_xref="GI:7301791"  
/translation="MISFLKLRNPQQLLLPSGVLSQWLKQKTMPIVYRCONPP  
KWYSAFGRNQKTGEMTADLDLNNWGRKSLMPLTGLACCAVEMHIAAPRYDM  
DGYVFRASPRQADLIIVAGTLTKMFAFKIYDQMPERWISMGSCANGGGYH  
YSJWVRGCDRIVPDIVYVGGPPTAEALMYGILOLQKKVKRMRTLOWMYR"  
join(46705..48163,54471..55572)  
/gene="msh"  
/note="Nucleotide sequence of the Celera sequence differs  
from the published sequence for this transcript."  
/product="CT2801"  
/db\_xref="FLYBASE:FBan0001897"  
/db\_xref="FLYBASE:FBgn0002849"  
/db\_xref="FLYBASE:FBgn0002849"  
/gene="msh"  
/note="CG1897"  
/map="99B4-99B5"  
/db\_xref="FLYBASE:FBan0001897"  
/db\_xref="FLYBASE:FBgn0002849"  
join(47104..48163,54471..54808)  
/gene="msh"



```

RESULT 10
AC009888      311043 bp      DNA      INV      04-SEP-1999
LOCUS      Drosophila melanogaster, chromosome 3R, region 99A7-99B5, BAC
DEFINITION      clones BACR48K01 and BACR48023, complete sequence.
AC009888      AC006093 AC007081
AC009888.1      GI:5822961
KEYWORDS      HYG.
SOURCE      fruit fly.
ORGANISM      Drosophila melanogaster
                Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
                Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
                Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE      1 (bases 1 to 311043)
AUTHORS      Celniker, S.E., Agbayani, A., Arcaina, T.T., Baxter, E., Blazej, R.G.,
                Butenhoff, C., Champe, M., Chavez, C., Chew, M., Ciesiolka, L.,
                Doyle, C.M., Farfan, D.E., Galle, R., George, R.A., Harris, N.L.,
                Hinkle, A., Hoskins, R.A., Houston, K.A., Hummasti, S.R., Karra, K.,
                Kearney, L., Kim, E., Lee, B., Lewis, S., Li, P., Ling, H.,
                Moshrefi, A.R., Moshrefi, M., Nixon, K., Pacleb, J.M., Park, S.,
                Pfeiffer, B., Poon, L., Richards, S., Sequeira, A., Sethi, H.,
                Swirskas, R.R., Wan, K.H., Webster, D., Woolley, P., Yang, S., Yee, M.,
                Zieran, L.L. and Rubin, G.M.
                Sequencing of Drosophila chromosome 3R, region 99A7-99B5
TITLE      Unpublished
JOURNAL      2 (bases 1 to 311043)
REFERENCE      Celniker, S.E., Agbayani, A., Arcaina, T.T., Baxter, E., Blazej, R.G.,
                Butenhoff, C., Champe, M., Chavez, C., Chew, M., Ciesiolka, L.,
                Doyle, C.M., Farfan, D.E., Galle, R., George, R.A., Harris, N.L.,
                Hinkle, A., Hoskins, R.A., Houston, K.A., Hummasti, S.R., Karra, K.,
                Kearney, L., Kim, E., Lee, B., Lewis, S., Li, P., Ling, H.,
                Moshrefi, A.R., Moshrefi, M., Nixon, K., Pacleb, J.M., Park, S.,
                Pfeiffer, B., Poon, L., Richards, S., Sequeira, A., Sethi, H.,
                Swirskas, R.R., Wan, K.H., Webster, D., Woolley, P., Yang, S., Yee, M.,
                Zieran, L.L. and Rubin, G.M.
                Direct Submission
                Submitted (04-SEP-1999) Berkeley Drosophila Genome Project, MS
                64-121, Lawrence Berkeley National Laboratory, One Cyclotron Road,
                Berkeley, CA 94720, US
COMMENT      On Sep 4, 1999 this sequence version replaced gi:5713324
                gi:5670685
                Sequence submitted by:
                Berkeley Drosophila Genome Project
                Lawrence Berkeley National Laboratory, MS 64-121
                Berkeley, CA 94720
                For further information about this sequence, including its location
                and relationship to other sequences, please visit our sequence
                archive Web site (http://www.fruitfly.org/sequence/) or send email
                to bdg@fruitfly.berkeley.edu.
FEATURES      Location/Qualifiers
                1..311043
                /organism="Drosophila melanogaster"
                /strain="y; cn bw sp"
                /db_xref="taxon:7227"
                /chromosome="3R"
                /map="99A7-99B5"
                /clone="BACs BACR48K01 (D488) and BACR48023 (D498)"
                /clone_lib="RPCI-98 (Roswell Park Cancer Institute
                Drosophila melanogaster BAC library, partial EcoRI in
                pBACe3.6)"
                /note="BACR48K01 (D488) and BACR48023 (D498) were
                sequenced as a project. BACR48K01 extends from bp 1 to bp
                183,934. BACR48023 extends from bp 142,325 to bp
                311,043."
BASE COUNT      87124 a 68661 c 68467 g 86791 t
ORIGIN
Query Match      37.4%; Score 308; DB 3; Length 311043;
Best Local Similarity      69.8%; Pred. No. 1.5e-41;
Matches 416; Conservative      0; Mismatches 180; Indels      0; Gaps      0;
Qy 127      ctgcgtccagcgtggcccgctgtgcaggcacagtggtccatcagcgtggccacc 186

```

```

Db 231214      CTGCTGGCAGTGGAGTGTGCAATCCAGTGCCTTCGTGGCCAGAGACCAATGCGGTG 231273
Qy 187      gatggcccaagcagcaccagcctgcccgtccaaagagccagagcctggtcccaaaccc 246
Db 231274      TACGATTACCGTTGCCAAAATCCGCCAAGTGGGGTACTCGGCCCTTCGAGGCAACCA 231333
Qy 247      agcagccggggcgagtgatgtggtggccaaagtgtgatactgaactggtcaactggccgcgcg 306
Db 231334      AAGACCTGGGGCAGTGGAGCTGTGCCAGTTGTGTGATGACTGCTAACTGGGAGACCAAG 231393
Qy 307      agttctctgtggcccatgaacttcgctggcctgtgctgctgctgctgctgctgctgctgctg 366
Db 231394      GGATCCCTCTGGCGCGTGACCTTTGGTCTCGCTGTGCGCCGTCGAAATGATGCACATT 231453
Qy 367      gacgacccgctacgacatggaccgctttgggtgtgtgttcttcgcgccagccgcgcg 426
Db 231454      CGGCTCTCGCTACGACATGGATCGTACGGTGTGTGTTCGGAGCATCTCTTCGCCAG 231513
Qy 427      tccagctcatgactgtggccggtgcacactcaccacaagaatggcccccagcgtcttcgaag 486
Db 231514      GCGACGCTGCTCATCGTGGCGCGGAACCTCGACCAACAAGATGCGCCGCTTTCGGAAG 231573
Qy 487      gtctacgacagatgcgcggagcgcgcgtacgtggttctccatggggagctgcgccacgga 546
Db 231574      ATCTACGACAGATGCCCGAGCGGAGATGGTTCATTTCCATGGCGAGTTCCGCCAATGTT 231633
Qy 547      ggagctactacactattctactcgtgtgtgtgagggctgcgacgcgcatctgtgccctgt 606
Db 231634      GCGGCTACTACCACTACTCTCTACTCGGTGTTCGGGCTGCGATGCGCATGTTCGGGTG 231693
Qy 607      gacatctacatccagctgcccacactacgcgcgagggccctgtctacggcatctcccgag 666
Db 231694      GACATCTACGTGCCGGATGTCCGCCACCGCGAGCGCTTAATGTACGGAATCTCTGCAG 231753
Qy 667      ctgcaggaggaagacaagcgggagcggagcgtgcagatctgtaccgcagtgatg 722
Db 231754      CTGCAGAGAAGAGTAAAGCGCATGAGGACCTCGAGATGGTACAGAAAGTAGAG 231809

RESULT 11
PDENQOVR      3739 bp      DNA      RCT      25-OCT-1993
LOCUS      Paracoccus denitrificans DNA repair protein (uvrA) gene, 3' end,
DEFINITION      and NADH dehydrogenase (NQO4, NQO5, NQO6, NQO7) genes, complete
                cds.
ACCESSION      M93015
VERSION      GI:409250
KEYWORDS      DNA repair protein; NADH dehydrogenase; NQO4 gene; NQO5 gene; NQO6
                gene; NQO7 gene; excision nuclease; ndhK gene; uvrA gene.
SOURCE      Paracoccus denitrificans
ORGANISM      Bacteria; Proteobacteria; alpha subdivision; Rhodobacter group;
                Paracoccus.
REFERENCE      1 (bases 1 to 3739)
AUTHORS      Xu, X., Matsuno-Yagi, A. and Yagi, T.
TITLE      Gene cluster of the energy-transducing NADH-quinone oxidoreductas
                of Paracoccus denitrificans: Characterization of four structural
                gene products
JOURNAL      Biochemistry 31, 6925-6932 (1992)
MEDLINE      92345253
REFERENCE      2 (bases 1 to 3739)
AUTHORS      Yagi, and T.
TITLE      Direct Submission
JOURNAL      Submitted (08-MAY-1992) Takao Yagi, The Scripps Research Institute
                Division of Biochemistry, SBR-15, La Jolla, CA 92037 USA
COMMENT      On Oct 22, 1993 this sequence version replaced gi:150618.
FEATURES      Location/Qualifiers
                1..3739
                /organism="Paracoccus denitrificans"
                /db_xref="taxon:266"
                1..333
                /gene="uvrA"
                source
                gene

```





```
* 23899 23978: gap of unknown length
* 23979 27079: contig of 3101 bp in length
* 27080 27159: gap of unknown length
* 27160 29602: contig of 2443 bp in length
* 29603 29682: gap of unknown length
* 29683 31919: contig of 2237 bp in length
* 31920 31999: gap of unknown length
* 32000 34542: contig of 2543 bp in length
* 34543 34622: gap of unknown length
* 34623 36445: contig of 1823 bp in length
* 36446 36525: gap of unknown length
* 36526 40018: contig of 3493 bp in length
* 40019 40098: gap of unknown length
* 40099 42062: contig of 1963 bp in length
* 42063 42141: gap of unknown length
* 42142 48164: contig of 6023 bp in length
* 48165 48244: gap of unknown length
* 48245 54296: contig of 6052 bp in length
* 54297 54376: gap of unknown length
* 54377 66627: contig of 12251 bp in length
* 66628 66707: gap of unknown length
* 66708 77734: contig of 11027 bp in length
* 77735 77815: gap of unknown length
* 77816 88534: contig of 10720 bp in length
* 88535 88614: gap of unknown length
* 88615 101020: contig of 12406 bp in length
* 101021 101100: gap of unknown length
* 101101 115583: contig of 14483 bp in length
* 115584 115663: gap of unknown length
* 115664 116236: contig of 573 bp in length
* 116237 116316: gap of unknown length
* 116317 116809: contig of 493 bp in length
* 116810 116889: gap of unknown length
* 116890 117343: contig of 454 bp in length
* 117344 117423: gap of unknown length
* 117424 118337: contig of 914 bp in length
* 118338 118417: gap of unknown length
* 118418 118910: contig of 493 bp in length
* 118911 118990: gap of unknown length
* 118991 119603: contig of 613 bp in length
* 119604 119683: gap of unknown length
* 119684 120333: contig of 650 bp in length
* 120334 120413: gap of unknown length
* 120414 120968: contig of 555 bp in length
* 120969 121048: gap of unknown length
* 121049 121685: contig of 637 bp in length
* 121686 121765: gap of unknown length
* 121766 122379: contig of 614 bp in length
* 122380 122459: gap of unknown length
* 122460 122949: contig of 490 bp in length
* 122950 123029: gap of unknown length
* 123030 123632: contig of 603 bp in length
* 123633 123712: gap of unknown length
* 123713 124146: contig of 434 bp in length
* 124147 124226: gap of unknown length
* 124227 124732: contig of 506 bp in length
* 124733 124812: gap of unknown length
* 124813 125420: contig of 608 bp in length
* 125421 125500: gap of unknown length
* 125501 126104: contig of 604 bp in length
* 126105 126184: gap of unknown length
* 126185 126663: contig of 479 bp in length
* 126664 126743: gap of unknown length
* 126744 127183: contig of 440 bp in length
* 127184 127263: gap of unknown length
* 127264 127803: contig of 540 bp in length
* 127804 127983: gap of unknown length
* 127984 128725: contig of 842 bp in length
* 128726 128805: gap of unknown length
* 128806 129167: contig of 362 bp in length
* 129168 129247: gap of unknown length
* 129248 129335: contig of 688 bp in length
* 129336 130015: gap of unknown length
```

```
* 130016 130560: contig of 545 bp in length
* 130561 130640: gap of unknown length
* 130641 131152: contig of 512 bp in length
* 131153 131232: gap of unknown length
* 131233 131826: contig of 594 bp in length
* 131827 131906: gap of unknown length
* 131907 132383: contig of 477 bp in length
* 132384 132463: gap of unknown length
* 132464 133016: contig of 553 bp in length
* 133017 133096: gap of unknown length
* 133097 133747: contig of 651 bp in length
* 133748 133827: gap of unknown length
* 133828 134276: contig of 449 bp in length
* 134277 134356: gap of unknown length
* 134357 134644: contig of 608 bp in length
* 134645 135044: gap of unknown length
* 135045 135383: contig of 339 bp in length
* 135384 135463: gap of unknown length
* 135464 136015: contig of 552 bp in length
* 136016 136095: gap of unknown length
* 136096 136547: contig of 452 bp in length
* 136548 136827: gap of unknown length
* 136828 137067: contig of 440 bp in length
* 137068 137147: gap of unknown length
* 137148 137600: contig of 453 bp in length
* 137601 137680: gap of unknown length
* 137681 138054: contig of 374 bp in length
* 138055 138134: gap of unknown length
* 138135 138352: contig of 218 bp in length
* 138353 138432: gap of unknown length
* 138433 138859: contig of 427 bp in length
* 138860 138939: gap of unknown length
* 138940 139149: contig of 210 bp in length
* 139150 139229: gap of unknown length
* 139230 139727: contig of 498 bp in length
* 139728 139807: gap of unknown length
```

Query Match 33.9%; Score 279.2; DB 2; Length 140685;  
Best Local Similarity 68.1%; Pred. No. 9e-37;  
Matches 403; Conservative 0; Mismatches 188; Indels 1; Gaps 1;

```
QY 133 tcacgctggccggcgtgtgcaggcacgaggtgtccatcagagcgtggccaccgatggc 192
Db 103599 TTGGCGCTGCCCGCTCTGAATTTGGCCCTGTGGCGCAACACAAACGCTTCCAGTTGC 103658

QY 193 ccaagcagcacccagcctgcctgccaaaggccagagcgtggctcccaaacccagcagc 252
Db 103659 CGAAGTGGCCCAAAATCTGCCCAAGAGGGGTACTTCTCGTTGGCCACCAACAGTCTC 103718

QY 253 cgggggc-gagtatgtgtggccaaagtggatgaactgtcaactgggcccgcggagttc 311
Db 103719 CGTTGGGAGTGGTCACTGGCCAGACTGGAGCATCTGCTCAACTGGGGTCCGAAGGCTC 103778

QY 312 tctgtggcccatgaacttcgacctgacctgctgcgctgagatgatgatgatgatgatgat 371
Db 103779 GATCTGGCCACTGACTTTGGCTTGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 103838

QY 372 accgcgtactacatggaccgctttggcggtgttcttcgcgcgcagccgcgcgcagtcgga 431
Db 103839 TCCGGGTTACGACATGGATCGATATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 103898

QY 432 cgtcatgatcgtggcggcgacactcaccacaagaatggcccccagcgttgcgaaggtcta 491
Db 103899 TGTCATCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 103958

QY 492 cgaccagatgccggagccgcgtacgtggtctccatggggagcgtgcgccaacgagagag 551
Db 103959 GCACCAAAATGCCGAGCCACACGTTGGTTCATCTCCATGGGAGCTGTGCCAACGCGCGG 104018

QY 552 ctactacacatttctactcgtggtgagggggtgcgaccgcacatcgtgcccgtggacat 611
Db 104019 CTACTACCATTAATCTGCTACTCCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 104078
```



* 45088	45167:	gap of unknown length	
* 45168	46233:	contig of 1356 bp in length	
* 46244	46803:	gap of unknown length	
* 46504	48559:	contig of 1956 bp in length	
* 48560	48639:	gap of unknown length	
* 48640	50589:	contig of 1950 bp in length	
* 50590	50669:	gap of unknown length	
* 50670	52607:	contig of 1938 bp in length	
* 52608	52887:	gap of unknown length	
* 52888	53881:	contig of 1194 bp in length	
* 53882	53961:	gap of unknown length	
* 53962	55649:	contig of 1688 bp in length	
* 55650	55729:	gap of unknown length	
* 55730	57438:	contig of 1709 bp in length	
* 57439	57518:	gap of unknown length	
* 57519	59734:	contig of 2216 bp in length	
* 59735	59814:	gap of unknown length	
* 59815	62621:	contig of 2807 bp in length	
* 62622	62701:	gap of unknown length	
* 62702	65239:	contig of 2538 bp in length	
* 65240	65319:	gap of unknown length	
* 65320	68682:	contig of 3363 bp in length	
* 68683	68762:	gap of unknown length	
* 68763	70859:	contig of 2097 bp in length	
* 70860	70939:	gap of unknown length	
* 70940	73115:	contig of 2176 bp in length	
* 73116	73195:	gap of unknown length	
* 73196	75979:	contig of 2784 bp in length	
* 75980	76059:	gap of unknown length	
* 76060	78750:	contig of 2891 bp in length	
* 78751	78830:	gap of unknown length	
* 78831	81983:	contig of 3153 bp in length	
* 81984	82063:	gap of unknown length	
* 82064	84253:	contig of 2190 bp in length	
* 84254	84333:	gap of unknown length	
* 84334	88351:	contig of 4018 bp in length	
* 88352	88431:	gap of unknown length	
* 88432	92028:	contig of 3597 bp in length	
* 92029	92108:	gap of unknown length	
* 92109	96970:	contig of 4862 bp in length	
* 96971	97050:	gap of unknown length	
* 102071	102071:	contig of 5021 bp in length	
* 102152	108106:	contig of 5955 bp in length	
* 108107	108186:	gap of unknown length	
* 108187	113640:	contig of 5454 bp in length	
* 113641	113721:	gap of unknown length	
* 113721	120443:	contig of 6723 bp in length	
* 120444	120523:	gap of unknown length	
* 120524	126527:	contig of 6004 bp in length	
* 126528	126607:	gap of unknown length	
* 126608	136785:	contig of 10178 bp in length	
* 136786	136865:	gap of unknown length	
* 137306	137305:	contig of 440 bp in length	
* 137306	137385:	gap of unknown length	
* 137386	137976:	contig of 591 bp in length	
* 137977	138056:	gap of unknown length	
* 138057	138524:	contig of 568 bp in length	
* 138525	138704:	gap of unknown length	
* 138705	139359:	contig of 655 bp in length	
* 139360	139439:	gap of unknown length	
* 139440	140066:	contig of 627 bp in length	
* 140067	140146:	gap of unknown length	
* 140147	140578:	contig of 532 bp in length	
* 140579	140758:	gap of unknown length	

Query Match 33.9%; Score 279.2; DB 2; Length 163378;  
Best Local Similarity 68.1%; Pred. NO. 8.5e-37;  
Matches 403; Conservative 0; Mismatches 188; Indels 1; Gaps 1;  
Qy 133 tccagctggccgcctgtgcagcacagatctccatccagagctgcccacccagatgcc 192  
Db 58610 TTGGGCGCTCCCGGCTCTGAATTGGCCCTGTGGCGCCAAACAAACGCTTCACGTTGC 58669

Qy 193	ccaagcagcaccacccagcctgcccctgcacaaaggccagagccgctggctccccccccccagcgc 252
Db 58670	CGAAGTGGCCCAAAATCTGCCCAAGAAGGGCTACTCTCGTTTCGGCAGCAACAGTCTCTC 58729
Qy 253	cggggc-gagtatgtgtgccaagtgtgatccctgtcaactgtgcccgcggagttc 311
Db 58730	CGTTCGGGAGTGTCTACTGGCAGACTGGACGATGTGTCTCACTGGGTGCGAAGGGTCTC 58789
Qy 312	tctgtgccccatgaccttcggcctggcctgctgcgcgcctggagatgacacatggcagc 371
Db 58790	GATCTGGCCACTGACTTTTCGGTTTGGCTGCTGTGCGCTGCGAATGATGACATCGCTGC 58849
Qy 372	acccctcagacatagacgcttggcgtgttccgcgcgcgcgcgcgcgcgcgcgcgcgcgc 431
Db 58850	TCCGCGTTACGATCGATATGTTGTGTTCGTGCTATCTCCAGCTGTCACGCTAGGCGCA 58909
Qy 432	cgtcatgatcgtggcggcacactcaccacaaagatggcccgagcgtctcgcaagttcta 491
Db 58910	TGTCATCATCTGCTGGCAGCTGACCAACAAATGGCACCGGCCCTGCGAAAGGTCTA 58969
Qy 492	cgaccagatgcgcggagccgcgtacgtgttcctccatggggagctgcgcaacggagagg 551
Db 58970	CGACCAATAGCCGACCGACGCTGATCTCCATGGGCACTGTGCCAACGGCGGG 59029
Qy 552	ctactaccactattcctactcgtgtgtgaggggctgcacccatcgtgcccgtggacat 611
Db 59030	CTACTACCATTTACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 59089
Qy 612	ctacatccacgctgcccaccctacggcggagcgcctgtctctacggcctctcgcagtcga 671
Db 59090	ATACGTACCGGGTTGTCGCCCAACCGCGAGCGGCTCATGTACGGCGTTCCTGACGCTGA 59149
Qy 672	gaggaagatcaagcggagcggagcgtgcagatcgtggtaccgcaggtagcgc 723
Db 59150	GAAGAAGTTAGCGGATGAACACGCTCCAGATGTGTTATAGCAAGTAAAGC 59201

RESULT 15  
AE003500/c  
LOCUS  
DEFINITION  
Drosophila melanogaster genomic scaffold 142000013386053 section 17  
of 30, complete sequence.  
ACCESSION  
AE003500 AE002593  
VERSION  
AE003500.2 GI:10728273  
SOURCE  
fruit fly.  
ORGANISM  
Drosophila melanogaster  
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;  
Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.  
1 (bases 1 to 327446)  
Adams,M.D., Celniker,S.E., Holt,R.A., Evans,C.A., Gocayne,J.D.,  
Amanatides,P.G., Scherer,S.E., Li,P.W., Hoskins,R.A., Gallie,R.F.,  
George,R.A., Lewis,S.E., Richards,S., Ashburner,M., Henderson,S.N.,  
Sutton,G.G., Wortman,J.R., Vandeil,M.D., Zhang,Q., Chen,L.X.,  
Brandon,R.C., Rogers,Y.H., Blazer,J.R.G., Champs,M., Pfeiffer,B.D.,  
Wan,K.H., Doyle,C., Baxter,E.G., Heit,G., Nelson,C.R., Gabor  
Miklos,G.L., Abrell,J.F., Agbayani,A., An,H.J.,  
Andrews-Pfannkuch,C., Baldwin,D., Ballew,R.M., Beeson,K.Y.,  
Baxendale,J., Bayraktarcglu,L., Beasley,E.M., Bolshakov,S., Borkova,D.,  
Botos,P.V., Berman,B.P., Bhandari,D., Bolshakov,S., Bortolier,P.,  
Botchan,M.R., Bouck,J., Brokstein,P., Brattier,P., Burtis,K.C.,  
Busam,D.A., Butler,H., Cadieu,E., Center,A., Chandra,I.,  
Cherry,J.M., Cawley,S., Dahlke,C., Davenport,L.B., Davies,P., de  
Pablos,B., Delcher,A., Deng,Z., Mays,A.D., Dew,I., Dietz,S.M.,  
Dodon,K., Doup,L.E., Downes,M., Dugan-Rocha,S., Dunkov,B.C.,  
Dunn,P., Durbin,K.J., Evangelista,C.C., Ferraz,C., Ferriera,S.,  
Fleischmann,W., Fosler,C., Gabriellian,A.E., Garg,N.S.,  
Gelbart,W.M., Glasser,K., Glodek,A., Gong,F., Gorrell,J.H., Gu,Z.,  
Guas,P., Harris,M., Harris,N.L., Harvey,D., Heiman,T.J.,  
Hernandez,J.R., Houck,J., Hostin,D., Houston,K.A., Howland,T.J.,  
Wei,M.H., Ibegwan,C., Jalali,M., Kalush,F., Karpen,G.H., Ke,Z.,



Kennison,J.A., Ketchum,K.A., Kimmel,B.E., Kodira,C.D., Kraft,C., Kravitz,S., Kulp,D., Lai,Z., Lasko,P., Lei,Y., Levitsky,A.A., Li,J., Li,Z., Liang,Y., Lin,X., Liu,X., Mattei,B., McIntosh,T.C., McLeod,M.P., McPherson,D., Merkulov,G., Milshina,N.V., Mobarry,C., Morris,J., Moshrefi,A., Mount,S.M., Moy,M., Murphy,B., Murphy,L., Muzny,D.M., Nelson,D.L., Nelson,D.R., Nelson,K.A., Nixon,K., Nusskern,D.R., Paclob,J.M., Palazolo,M., Pittman,G.S., Pan,S., Pollard,J., Puri,V., Reese,M.G., Reinert,K., Remington,K., Saunders,R.D., Scheeler,F., Shen,H., Shue,B.C., Siden-Kiamos,I., Simpson,M., Skupski,M.P., Smith,T., Spier,E., Spradling,A.C., Stapleton,M., Strong,R., Sun,E., Swirskas,R., Tector,C., Turner,R., Venter,E., Wang,A.H., Wang,X., Wang,Z.Y., Wasserman,D.A., Weinstein,G.M., Weissenbach,J., Williams,S.M., Woodage,T., Worley,K.C., Wu,D., Yang,S., Yao,Q.A., Ye,J., Yeh,R.F., Zaveri,J.S., Zhan,M., Zhang,G., Zhao,Q., Zheng,L., Zheng,X.H., Zhong,F.N., Zhong,W., Zhou,X., Zhu,S., Zhu,X., Smith,H.O., Gibbs,R.A., Myers,E.W., Rubin,G.M., and Venter,J.C.

The genome sequence of *Drosophila melanogaster*  
 Science 287 (5461), 2185-2195 (2000)

20196006

2 (bases 1 to 327446)

Adams,M.D., Celniker,S.E., Gibbs,R.A., Rubin,G.M. and Venter,C.J.  
 Direct Submission  
 Submitted (21-MAR-2000) Celera Genomics, 45 West Gude Drive,  
 Rockville, MD, USA

On Oct 9, 2000 this sequence version replaced gi:7293108.

FEATURES  
 Location/Qualifiers  
 1..327446  
 /organism="Drosophila melanogaster"  
 /db\_xref="taxon:7227"  
 /chromosome="X"  
 complement(join(772..903,968..1504,1572..2077,2153..2821,  
 4145..>4380))  
 /gene="CG9281"  
 /product="Cr26402"  
 /db\_xref="FLYBASE:FBan0009281"  
 /db\_xref="FLYBASE:FBgn0030672"  
 complement(join(772..903,968..1504,1572..2077,2153..2813))  
 /gene="CG9281"  
 /note="CG9281 gene product"  
 /codon\_start=1  
 /db\_xref="FLYBASE:FBan0009281"  
 /db\_xref="FLYBASE:FBgn0030672"  
 /protein\_id="AAF48493.1"  
 /db\_xref="GI:7293109"  
 /translation="MSSAAKREAQKKDKARKNKIQIPSTKKANGVPERLTEEKL  
 CALEEARISAEARCTGSLVHPSRDVKIANFTIPFGSELLQDTMLELNCGRY  
 GLTLCGCKSSLLVGLGVEVPVPHIDIFHLREIPASSKSLQCVMEVDEIRKL  
 EKAAELAMSEEDAOEQLIDYVERLDDMSADLAEVKARIILHGLGFQKAMQOKAKD  
 FSGWRNRILARALFVKPHLLLDLDEPTNHLDDACVLEELKYKYLILSHSQD  
 FLNGVCNIIHLTKKLLKYTGNYAEVFRVRELLLENQKQINWQDDQISHKNTIAR  
 FGHGAKLAQSKKTKLAKVAQGLTEKVDKLVNFPFSCGKVPVPPVIMQNVN  
 FRYNDEPTWYKLEFGIDIDTLALVGNAGKSTLLKLYGDLVPTSGMRKNSHL  
 RIARYHOLLHLLDASPLEYAMRAFDPVQKEEMRKIGRYGLTGRCQVCPIRQLS  
 DGQRCRVFVAALAVQVPHLLDEPTNHLDMETIDALADAINDFGGMVLVSHDFRLI  
 NQVARELWCEKTYKWKGGILDYKDKHLNKTITSNEKKAKAGK"  
 join(<4963..5295,5348..>5848)  
 /gene="CG15601"  
 /product="CT35724"  
 /db\_xref="FLYBASE:FBan0015601"  
 /db\_xref="FLYBASE:FBgn0030673"  
 /evidence=not\_experimental  
 <4963..>5848  
 /gene="CG15601"  
 /map="13E14-13E14"  
 /db\_xref="FLYBASE:FBan0015601"  
 /db\_xref="FLYBASE:FBgn0030673"  
 /evidence=not\_experimental

join(4963..5295,5348..5848)  
 /gene="CG15601"  
 /note="CG15601 gene product"  
 /codon\_start=1  
 /db\_xref="FLYBASE:FBan0015601"  
 /db\_xref="FLYBASE:FBgn0030673"  
 /evidence=not\_experimental  
 /protein\_id="AAF48494.1"  
 /db\_xref="GI:7293110"  
 /translation="MSKUSDTKIPITELEAYRROPCLYNLLSDSKYKRVSRREAYGA  
 IYSLKIPQTVSDIKLKIKSVRTYVSKELRIWMREKEIGRTYEPKLFWRADSFRLR  
 SVSLCKRQKSNSSAQLTTIKSDTSKLCTAAADITMSDEDAEEDAEVNEPE  
 ECPLEBSRTASICKDDSTLCADQQEHSYQSGSSQQLPHTMAQKRSYITSLDS  
 AGEDDLIFGQSIASQLRTIPDSYSKSAKLRIQOVLEAETGQGFSTEVASTOLQNT  
 F"  
 join(<7954..8220,8556..8795,8865..9086,9145..9416,  
 9497..9993,10131..10515,10567..15238,15758..18218,  
 18282..18755,18814..19154,19217..20287,20378..20513,  
 20573..22032,22060..23197,23270..23671,23739..23927,  
 23992..24313,24377..24542,24617..24807,24885..>24987)  
 /gene="CG8184"  
 /product="Cr24358"  
 /db\_xref="FLYBASE:FBan0008184"  
 /db\_xref="FLYBASE:FBgn0030674"  
 <7954..>24987  
 /gene="CG8184"  
 /map="13E14-13E16"  
 /db\_xref="FLYBASE:FBan0008184"  
 /db\_xref="FLYBASE:FBgn0030674"  
 join(7954..8220,8556..8795,8865..9086,9145..9416,  
 9497..9993,10131..10515,10567..15238,15758..18218,  
 18282..18755,18814..19154,19217..20287,20378..20513,  
 20573..22032,22060..23197,23270..23671,23739..23927,  
 23992..24313,24377..24542,24617..24807,24885..24987)  
 /gene="CG8184"  
 /note="CG8184 gene product"  
 /codon\_start=1  
 /db\_xref="FLYBASE:FBan0008184"  
 /db\_xref="FLYBASE:FBgn0030674"  
 /protein\_id="AAF48495.1"  
 /db\_xref="GI:7293111"  
 /translation="MKVDVAKLKKGSSLTABRELIDELTKRRTRAELEFLDTHVHV  
 WIYKGELEHWEILDRFEXLLEASRHYNGNEFVLOCTNFQETDITLLHLNFTT  
 LLIEHSFRLYNSIEHLIVLLSSQNDIVLAVLLLYMFSKSNFIPLRPEKKELL  
 IVKLFSAERWDPNYGLSKDCGCEKLEFLYLQICIDYDVERHQAQLEIPDMMDL  
 CHTASAPVTKIAGQISKPSEAIKMRIAHVRVLSGFNNYKLRQLQVQARQVAVSIL  
 IYSNALQNDTKVLYAGFGECELDKEDVHVEIRAAVLTLSMLHDFDNWPS  
 SYWGCTALVKSCMQSLICVSWPGVDLEHITFVTRAVRVIDLTINIDIAFHQNG  
 LVNFIORLEIKSCAKALADIGITLKESTLRDCHKLDTSLDHVEDVEDGGENASV  
 TSNASPRDNDSDGATGAGSIVYRAYHEHIGAGSGAGLLDIGARSLSGSSGTS  
 GMGASINRDGSAARPAIMYPNTSYIARPLAEKRAELSSQLPSLAKPKPCVQRA  
 ALLNNRNIMETSLQSLRHITANAEEYGPSLFLATDVTVTVVNEPSLLSSLODLG  
 ITSVMKALLQKQVPATREVIGSLPNVFSALCINERGLFEFLSYDPFKVLKVLSPD  
 YLVAMRRRSRSLDGTATNLGNAMDLIKHHPYLRADATEAIRLLNELVRLGSDPS  
 FICWRANKSSGSGSGHGGHVAVSTPSPMVAVAGTSASSVQGAADTNDSSGD  
 DDDDDMSASQQQQPTPGCGGSPSTPQCGVVGSGAGATGTPAAASQAVKV  
 VTPPERAFLIDITLNNWKFIEAFISNPGDHCFREFVLHGLAPILQLLSLNLPLV  
 DSVSTQSAVANYKLSQAQETKVLDAQIADIVAQKLKLIKHTFIPFGGSVLL  
 AELVQCORLDGFANAEYTPILHNSFVHGYVVMVHLCRNASNMRTILLKRWGNR  
 ETGVQLQQLVQLYISLVNMTILLNCDEPTQHPDLIDIAQMSAAGTADPLT  
 EAEISRLERAEEFRKTYTPEOFRYIKSLAASSRLGRALAEGLTFLVLSVGSPT  
 RQRINDLISNITKVTPEARILISITLYNGFSHESILPKPVPKLTFLICSVG  
 FTGPMFLDDRRSAFLNISOQCAENGLKAFEFMYWALSIDNVSOKEFDQWASMEDL  
 MQSHEDDKDCPEGTGEFDLWQLLEKAVNPRMLDSTYMSPALNYPDVSVPREP  
 FYLIHQLAMQALRRIRWARRPINYPNGSPSMTMILIKHLKSHQTLDRTHRMKE  
 IKFENLYIRSTDDGLNADHIKTLTDMGFHYHVIEALRTNASLEADYLLNPEASS  
 LSTTGQSSAPPAPPPPPASMTDIDVDPADGESQSKSTSPNSYDKHLKMLP  
 SLIFDRECEAISRAREWEAIPDVTNSAADLAVLYRRHNLKVPYFVNEVRNIQ  
 WYDFQLPFPATGSKASLEECITGMTATFLFVRLKTLSTLLPENYSDLRPLVE  
 ATTAQDAWLSLVKLLDCSMQWMEQSTENTGOORATVPRWQNLVDLDAIDLSHI  
 LQRKANNRVSSDTHWRDASTGKNWAYSEANNELRNAYFAGERWLYINIGRQCTV  
 SLNCTQVSEASGTHRPVCPALKLSAITSNNPVALHKVHEHLNVRVTPDGGCTV

HSDLLISLRQLVNETDGOPOQPOQSENSANPSEATSEDCPAAANSQSPFPATMTTRSKTR									
QKSAARAAANKSQSGSPFPARTQKILINEEQVGLSKDFCDRIIGSIVDILQPSH									
QMLNETKLSLRRLRNADQVIRRGVERMLLHQACSFMGFTYANILIRH									
VLEAPVLOEAMERVGLNRAGIYVGHROLIYVITQVSAVSRNFQIFVAAAKEMLK									
GEYLINTSTLLADARVYKSKFGQPTAAAGATPHNOTPTPPAAKDELKDDPQV									
QSSVALKELHLLGVQVHYGTSLSDRQGAELQDPDQPGTAQCGCHDDMTVL									
ICSLDCYQOPGVQVHAEELHAAVLAQPSAKKHTRLQVLMISLPNLIESPMLY									
DKVHHRNMVRLVLRQGVMTYLPKVAQYKDLNSHNITLSLAAILRPMELLRFVSST									
TTLGSKRILFGSGPGPNASGNNTAGSSGNGSPGNTYGTIINRRLYPLAAAT									
GFGSARNSAMGEHVLRLDIRVLSDRRHAFEFINSDDMADVSDAAAPGSGVGG									
GGSGAGVGGVSGSVSSNSAAVIDVEERSNGDNDPPANVSSAGSAGSSTSG									
GPRVVFVDFDQLEGLRLASRGAGSGSGVSGVSSGQNGSNTASSTN									
SGNSGICNLRLNTPVDVSGSGSGSNAGVYNDNGISGGSGSGCASSTGT									
APGGARDLNASLSDSDSTENDERNEDDEDEDDADDEDDVDEHSDTDE									
PESROFIEVFPHIVEPESSETASNDADVNDNDDDDDDDDDDDDDDDDVREVO									
SVMDAEINRDIALLAYSSNNRPRSAVNASGDSANRTNEDDEDDDDDDADP									
DGDVDVDFISPSAQAQEAZYLYRLGSLMPTPPVGRVWNSSIAQPLDIDVNGNA									
SGSNAGATGSSGSGRSTGGTTASIGATASAAALRPSGSMWAPDNFIRISGDS									
SNGSPALGGNANGDMNVFTTPSMLERAGAGPGSQTNSELLNEQPTGAGGGGG									
SASGVGVGANGTGHNPATLCTNGVRRRLYLDOQYCVCMPTNQNPTTEQMETF									
TQDALHWWLEBATLIDMESQTDACLYAAHLLPHLIKELKAAHQKQEEAARQSTTP									
Query Match 33.9%; Score 279.2; DB 3; Length 327445;									
Best Local Similarity 68.1%; Pred. No. 6.5e-37;									
Matches 403; Conservative 0; Mismatches 188; Indels 1; Gaps 1;									
QY 133	tcacgctggtggccggtgtgcagcgaggtgtccatcacagcgtggccaccgatggc	192							
Db 276717	TTCCGGCGTGCCTGCTCAATTTGGCCCTGGTGGCCCAACAACACGCTTCACGTGC	276658							
QY 193	ceaagcagcaccacgctccctgcacaaagccagagcgcgtggtcccaaacccagcgc	252							
Db 276657	CGAAGTGGCCCCAAAATCTCCCAAGAGGGCTACTCTCCGTTCGGCACCAACAGTCTC	276598							
QY 253	cggggc-gagtatgtgtgccaagctggtgacccctgacccgagccgcgcgcgcgcgc	311							
Db 276597	CERTGCGAGTGTCTACTGCCAGACTGGACGATCTGCTCAACTGGGTGCGAAGGCTC	276538							
QY 312	tcgtgtgcccacgaccccttggcctggtgctgctgctgctgctgctgctgctgctg	371							
Db 276537	GATCTGGCCACTGACTTTGGCTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	276478							
QY 372	accccgctacgacatggaacgtttggcgtgtcttcgcgcgcgcgcgcgcgcgcgcgc	431							
Db 276477	TCCGCGTTACGACATGGATCGATATGGTGTGTGCTGCTGCTGCTGCTGCTGCTG	276418							
QY 432	cgctacgctggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtg	491							
Db 276417	TGTATCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	276358							
QY 492	cgcacagatgc	551							
Db 276357	CGACCAATGCGCGGACCGACGTTGGTGTATCTCCATGGGCGAGCTGTGCCAACGCG	276298							
QY 552	ctactaccactattcctactcgtggtggtggtggtggtggtggtggtggtggtggtg	611							
Db 276297	CTACTACCATTAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	276238							
QY 612	ctacatccagagctccacacccacccacccacccacccacccacccacccacccac	671							
Db 276237	ATACGTACCCGGTTGCCCCCAACCGCGAGCGCTCATGTACGCGCTTTTGACGCTG	276178							
QY 672	gaggaagatcaagcgggagcgagcgctgagctgagctgagctgagctgagctgagctg	723							
Db 276177	GAAGAAGGTTAAGCGATGAAGACGCTCCAGATGTGGTATAGGAAGTAAGGC	276126							

XX	AAV58509;	
AC		
XX	08-DEC-1998 (first entry)	
DT		
XX		
XX	3' fragment of prostate tumour specific gene J1-21.	
DE		
XX	Prostate tumour specific gene; human; prostate cancer; detection;	
KW	therapy; ss.	
KW		
XX		
OS	Homo sapiens.	
XX		
XX	WO9837418-A2.	
PN		
XX		
PD	27-AUG-1998.	
XX		
PF	25-FEB-1998; 98WO-US03690.	
XX		
XX	09-FEB-1998; 98US-0904809.	
PR	25-FEB-1997; 97US-0806596.	
PR	01-AUG-1997; 97US-0904809.	
XX		
XX	(CORI-) CORIXA CORP.	
PA		
XX		
PI	Dillon DC, Xu J;	
XX		
XX	WPI; 1998-480805/41.	
XX		
XX	Novel human prostate specific tumour protein and fragments - useful	
PT	for detecting and treating prostate cancers	
XX		
XX	Claim 1; Page 50; 14ipp; English.	
PS		
XX		
CC	This sequence represents a human prostate tumour specific gene, and can	
CC	be used in the method of the invention. The method is for detecting	
CC	prostate cancer comprises contacting a biological sample with an agent	
CC	able to bind an immunogenic portion of a prostate protein (such as	
CC	encoded by this sequence). An antibody which binds to an immunogenic	
CC	portion of the prostate protein, and the method can be used to detect,	
CC	monitor progression of, or treat prostate cancers. The antibody may	
CC	also be conjugated to a therapeutic agent for use in therapy of prostate	
CC	cancers.	
XX		
SQ	Sequence 789 BP; 129 A; 240 C; 159 G; 195 T; 66 other;	

Qy	491	acgaccagatgcgcggagccgcgtactcgtggtctccatgaggagctgcgcgaacgcgagag	550
Db	381	AAGACCAGATGCCGGAGCCCGGANAAGTGGTTTCCANGGGAG-TGNGCCCAACGGNGNG	323
Qy	551	gctactaccactatctctactcgtggtgagggctgcacacgcacgcgtgcccgtagaca	610
Db	322	GGTAANACCACCTATTCCTAATNGTGGTGAGGGGNTGGACCGCAITGTGCCCGTGGAA	263
Qy	611	tctacatccacaggctgcgcacactacgcgcgagccctgctctctacgcacactcgtgcagctgc	670
Db	262	TTNAAATCCCAAGGTGCCACCNANGCCGAGCCCTGTTTANGGCATCTCTGCAGCTGC	203
Qy	671	agaggaagatcaagcgggagcgaggctgcagatctggtaccgcaggttagcgcgcgcgc	730
Db	202	AGAGGAANATCAAGCGGAGCGGAGCTGCANATTTGGTACCGACAGTAGGCGCGCCGCC	143
Qy	731	gcgcgcgcgcgagccctgcgcgcctcgtctgtccacagcctgctgtgtcccgtagaggtgtgc	790
Db	142	GCCGCCGCCGAGCGCTTTCGCCCTCTGTCTCCACAGCCTGNTTGTGTCCCGNAGAGGTGTGC	83
Qy	791	aataaacctgccttcggcgcaaaaaaaaaaaaaa	824
Db	82	AATAAACCTGCCTTNGNAAAAAAAAAAAAAAAAA	49

RESULT	6	
AAA06272/c		
ID	AAA06272	standard; cDNA; 789 BP.
XX	AC	
XX	AA06272;	
XX		
XX	DT	13-JUN-2000 (first entry)
XX	DE	Human immunogenic prostate tumour protein cDNA sequence SEQ ID NO:32.
XX	KW	Human; prostate cancer; diagnosis; tumour; gene therapy; detection;
XX	KW	immunogenic; cytostatic; vaccine; ss.
XX		
XX	OS	Homo sapiens.
XX	OS	
XX	PN	WO200004149-A2.
XX	PD	
XX	PD	27-JAN-2000.
XX		
XX	PF	14-JUL-1999; 99WO-US15838.
XX		
XX	PR	14-JUL-1998; 98US-0115453.
XX	PR	14-JUL-1998; 98US-0116134.
XX	PR	23-SEP-1998; 98US-0159812.
XX	PR	23-SEP-1998; 98US-0159822.
XX	PR	15-JAN-1999; 99US-0232149.
XX	PR	15-JAN-1999; 99US-0232880.
XX	PR	09-APR-1999; 99US-0288946.
XX		
XX	PA	(CORI-) CORIXA CORP.
XX		
XX	PI	Dillon DC, Harlocker SL, Yugu J, Xu J, Mitcham JL;
XX	PI	WPI; 2000-171268/15.
XX	DR	
XX		
XX	PT	New polypeptide useful for treating and diagnosing prostate cancer
XX	PT	comprises an immunogenic portion of prostate tumor protein -
XX		
XX	PS	Claim 50; Page 110; 263pp; English.

xx The present invention describes isolated polypeptides, comprising an  
CC immunogenic portion of a prostate tumour protein (ptp). The polypeptides  
CC and polynucleotides encoding them have cytotoxic activity and can be  
CC used in vaccines and in gene therapy. The polypeptides and  
CC polynucleotides encoding them, antigen presenting cells which express  
CC the polypeptides, antibodies against the polypeptides and vaccines  
CC comprising them can be used for inhibiting the development of prostate





(CORI-) CORIXA CORP.

Xu J, Dillon DC, Mitcham JL, Harlocker SL, Jiang Y, Reed SG; Kaos MD, Retter MW, Stolk JA, Day CH, Skeiky YAW, Wang A; WPI; 2001-308785/32.

Isolated polypeptide comprising at least an immunogenic portion of a prostate-specific protein, useful in the diagnosis and therapy of prostate cancer -

Claim 31; Page 140; 325pp; English.

The present invention describes an isolated polypeptide (PI) comprising at least an immunogenic portion of a prostate-specific protein, or its variant. Also described are polynucleotides (NI) encoding (PI). (PI) and (NI) have cytostatic activity and can be used in vaccine production. The polypeptides, nucleic acids and antibodies from the present invention are useful in the diagnosis and therapy of prostate cancer. Prostate specific genes P704P, P712P, P774P, P775P and P305D are located in a genomic region on chromosome 22q1.2 known as the Cat Eye Syndrome region. Prostate specific antigen (PSA) P501S was located on chromosome 1. AAH84571 to AAH85143 and AAG39000 to AAG39077 represent polynucleotide and polypeptide sequences used in the exemplification of the present invention.

Sequence 789 BP; 129 A; 240 C; 159 G; 195 T; 66 other;

Query Match 46.7%; Score 384.6; DB 22; Length 789;  
Best Local Similarity 75.9%; Pred. No. 2.1e-75;  
Matches 527; Conservative 0; Mismatches 154; Indels 13;

141	QY	gggcccgcgtctgcaggacagagtgctccatcatagagtggtgccaccgatggcccaagacg	200
739	Db	GGGGGATGGTNNNAGANGNGTTTTTCNTTAGNNNGGCCACCGGGGGCCCGAAGG	680
201	QY	cacccagcctccctgccaaagccagagccgtgctctccaaacccagccagcggggcca	260
679	Db	GGANCCANTTGCNCNCCAAAGNNAGAGCGTGGGTTCCCAAGCCAGNAANGCGGGGA	620
261	QY	gtatgtgtggccaaagctggatgacctcgt---caactggccgcgcgcgagttctctctgtg	317
619	Db	GCATTTGGGGGCCAAGGGGATGACTTTCTCAAACTGGNCCCCCGGAGNTTTTTCGN	560
318	QY	g-----cccatgaccttgcgcctgcctgtgcgcgtgagatgatcacatggcagc-a	372
559	Db	GGGCCNCAAGAAACATTTTGGGTTCNCCTGTGNGCCCGGGAGGTGATGNATNGCAGCA	500
373	QY	cccgcctacgacatgagccgctttgcgtgg--tcttcgcgcgcagccgcgcagctcgc	430
499	Db	CCCCGNNAGACATGNNCCGGTTTGGGGNGNTTTTGGGGGCCAGCCGCGCAGTCGTG	440
431	QY	acgtcatgatgtggccgcgcacactcaccacaagaatggcccccgcgtctcgaaagtct	490
439	Db	A-GTNATGATTGGCNGGNANANTAACCAAAAGAGGGCCCCAGNG-TTTGCAAGGTTT	382
491	QY	acgaccagatgcggagacgcgcgaactgcttcacatggggagctgcgcacgcggagag	550
381	Db	AAGACCAAGTCCGGAGCCGCGANAAAGTGGTTTCCANGGGAG-TGNGCCACAGCGNGNG	323
551	QY	gctactacacattctctactcgttggtgaggggctgcgcacccatgctgcgcgtgaca	610
322	Db	GGTAANACCACTATTCTAATNGGTGGTGAGGGGNTGGCCCGCATTTGTCCCCGTGGAAA	263
611	QY	tctacatccaggctgcaccaactcagccgcagggccctcgtctcagcactcctgcagctgc	670
262	Db	TTNAAATCCAGGGTGCACCACCNANGCCGAGGCCCTTTTANGGCATCTTCACCTGC	203
671	QY	agaggagatcaagcgggagcagagctgcagatctcgtgtaccgcagtagcgcgcgcgc	730
202	Db	AGAGGAANATCAAGCGGAGCGGAGGCTGCANATTGTGTACCGCAGTAGGCGCGCGCG	143

[illegible]

## RESULT 10

AAH02453/c  
ID AAH02453 standard: cDNA: 789 BP.

XX AH02453.

14-JUN-2001 (first entry)

DE Prostate tumour antigen determined 3' cDNA sequence for J1-21.

Human; prostate tumour antigen; prostate tumour; therapy; diagnosis

XX OS Homo sapiens.  
XX OS Homo sapiens, cytotaxic, vaccine, ss.  
XX OS Homo sapiens, cytotaxic, vaccine, ss.

AA  
PN  
WO200125272-AAA  
PD 12-APR-2001.

04-OCT-2000:

PR 04-OCT-1999; 99US-0157455.

PA (CORI-) CORIXA CORP.

PI Xu J, Skeiky YAW,

DR WPI; 2001-245062/25.

PT Prostate specific pr

XX

XXI

at least an immunogenic portion of a prostate tumour antigen protein or its variant. (I) have cytostatic activity and can be used in vaccine production. (II), prostate tumour antigen polynucleotides, an antigen presenting cell (APC e.g. a dendritic cell) that expresses (I), and a pharmaceutical composition containing (I) are useful for inhibiting the development of cancer in a patient. Antibodies specific for prostate specific proteins and oligonucleotides that hybridise to a polynucleotide that encodes a prostate specific protein are useful for detecting the presence or absence of a cancer or monitoring the progression the progression of a cancer, especially prostate cancer. AAH02422 to AAH2872, AAB74798 to AAB74821 and AAB74830 are sequences used in the exemplification of the present invention.

Sequence 789 BP; 129 A; 240 C; 159 G; 195 T; 66 other;

Query Match 46.7%; Score 384.6; DB 22; Length 789;  
Best Local Similarity 75.9%; Pred. NO. 2.1e-75;  
Matches 527; Conservative 0; Mismatches 154; Indels 13;

QY	141	gggcccgcctgcaggcacacgagttccatcagacagctggccaccgatggcccaagcag	200
Db	739	GGGGATGTTNNAGANGNGTTTCTNTTAgANNNGGCCACCGGGGCCCGAAGGG	680
QY	201	caccagcctgcctgcacaaaggccagccctggctcccaaaccccgacgcgcgggcga	260
Db	679	GGANCCACTTGCCTNCCCAAGNAGAGCGGGTTCCTCAACCCAGNAANCGGGGGA	620
QY	261	gtatgtgtggccaaagtgcgatgacctgt---caactggggccgcgcggagttctctgtg	317

RESULT 11





KW metabolic pathway; promoter; termination sequence; ss.

XX Arabidopsis thaliana.

XX PN EP1033405-A2.

XX PD 06-SEP-2000.

XX PF 25-FEB-2000; 2000EP-0301439.

XX XX 25-FEB-1999; 99US-0121825.

PR 05-MAR-1999; 99US-0123180.

PR 09-MAR-1999; 99US-0123548.

PR 23-MAR-1999; 99US-0125788.

PR 25-MAR-1999; 99US-0128264.

PR 29-MAR-1999; 99US-0128785.

PR 01-APR-1999; 99US-0127462.

PR 06-APR-1999; 99US-0128234.

PR 08-APR-1999; 99US-0128714.

PR 16-APR-1999; 99US-0129845.

PR 19-APR-1999; 99US-0130077.

PR 21-APR-1999; 99US-0130449.

PR 23-APR-1999; 99US-0130510.

PR 28-APR-1999; 99US-0130891.

PR 30-APR-1999; 99US-0131449.

PR 30-APR-1999; 99US-0132048.

PR 04-MAY-1999; 99US-0132407.

PR 05-MAY-1999; 99US-0132484.

PR 06-MAY-1999; 99US-0132485.

PR 06-MAY-1999; 99US-0132486.

PR 07-MAY-1999; 99US-0132487.

PR 11-MAY-1999; 99US-0132863.

PR 14-MAY-1999; 99US-0134256.

PR 14-MAY-1999; 99US-0134218.

PR 14-MAY-1999; 99US-0134219.

PR 14-MAY-1999; 99US-0134221.

PR 14-MAY-1999; 99US-0134370.

PR 19-MAY-1999; 99US-0134768.

PR 20-MAY-1999; 99US-0134941.

PR 20-MAY-1999; 99US-0135124.

PR 21-MAY-1999; 99US-0135353.

PR 21-MAY-1999; 99US-0135629.

PR 25-MAY-1999; 99US-0136021.

PR 27-MAY-1999; 99US-0136392.

PR 28-MAY-1999; 99US-0136782.

PR 01-JUN-1999; 99US-0137222.

PR 03-JUN-1999; 99US-0137528.

PR 04-JUN-1999; 99US-0137502.

PR 06-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0138817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149923.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.



PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139453.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144684.  
PR 21-JUL-1999; 99US-0144684.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145313.  
PR 27-JUL-1999; 99US-0145318.  
PR 27-JUL-1999; 99US-0145319.  
PR 28-JUL-1999; 99US-0145351.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 05-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.

PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155859.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 14-OCT-1999; 99US-0159638.  
PR 18-OCT-1999; 99US-0159584.  
PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160768.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 21-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160980.  
PR 22-OCT-1999; 99US-0160981.  
PR 22-OCT-1999; 99US-0160989.  
PR 25-OCT-1999; 99US-0161404.  
PR 25-OCT-1999; 99US-0161405.  
PR 25-OCT-1999; 99US-0161406.  
PR 26-OCT-1999; 99US-0161359.  
PR 26-OCT-1999; 99US-0161360.  
PR 26-OCT-1999; 99US-0161361.  
PR 28-OCT-1999; 99US-0161820.  
PR 28-OCT-1999; 99US-0161992.  
PR 28-OCT-1999; 99US-0161993.  
PR 29-OCT-1999; 99US-0162142.

Query Match 30.6%; Score 251.8; DB 21; Length 976;  
Best Local Similarity 69.8%; Pred. No. 2.4e-46;  
Matches 340; Conservative 0; Mismatches 147; Indels 0; Gaps 0;





KW Neisseria meningitidis; Neisseria gonorrhoeae; genome; immunogenic;  
 KW antigen; vaccine; diagnosis; infection; antibacterial; identification;  
 KW Meningococcus B; MenB; ds.  
 XX  
 XX Neisseria meningitidis.  
 XX  
 XX WO200022430-A2.  
 XX  
 XX 20-APR-2000.  
 XX  
 XX 08-OCT-1999; 99WO-US233573.  
 XX  
 XX 09-OCT-1998; 98US-0103794.  
 PR 30-APR-1999; 99US-0132068.  
 XX  
 XX (CHIR ) CHIRON CORP.  
 PA  
 XX Frazer CM, Hickey E, Peterson J, Tettelin H, Venter JC;  
 PI Masignani V, Galeotti C, Mora M, Ratti G, Scarselli M, Scarlato V;  
 PI Rappuoli R, Pizza M;  
 XX  
 DR WPI; 2000-318079/27.

Isolated nucleotide sequences of Neisseria meningitidis which can be used in the diagnosis and treatment of N. meningitidis infection and other Neisserial infections, for example, N.gonorrhoea -

Claim 7; Page 406-428; 1760pp; English.

XX The present invention describes methods of obtaining immunogenic  
 CC proteins from Neisseria genomic sequences. AAA81453 to AAA82414  
 CC represent specifically claimed Neisseria meningitidis genomic DNA  
 CC sequences; AAA81260 to AAA81303 and AAB25620 to AAB25663 represent  
 CC Neisseria DNA sequences and their corresponding proteins; AAA81254 to  
 CC AAA81259 and AAA81304 to AAA81321 represent PCR primers used in the  
 CC isolation of Neisseria meningitidis DNA sequences; and AAA81322 to  
 CC AAA81452 represent Neisseria meningitidis MenB polynucleotide ORF  
 CC sequences, which are all used in the exemplification of the present  
 CC invention. The nucleic acid sequences, protein sequences, and antibodies  
 CC against them, can be used in the manufacture of a composition. The  
 CC composition can be used as a medicament (or in the manufacture of a  
 CC medicament) for treating, preventing or diagnosing infection due to  
 CC Neisserial bacteria. For example, some of the identified proteins could  
 CC be components of vaccines against Meningococcus B; against all serotypes;  
 CC and/or against all pathogenic Neissariae. Identification of sequences  
 CC from the bacterium will also facilitate production of biological probes,  
 CC particularly organism-specific probes. Attempts to make efficacious  
 CC Meningococcus B vaccines have failed mainly due to antigen tolerance.  
 CC Multivalent vaccines have also been tried but none have successfully  
 CC overcome antigenic variability. The provision of further, complete  
 CC sequences may provide an opportunity to identify secreted or surface  
 CC exposed proteins that may be presumed targets for the immune system and  
 CC which are not antigenically variable or at least more conserved than  
 CC other more variable regions.

XX Sequence 72750 BP; 17518 A; 19945 C; 18810 G; 16477 T; 0 other;

Query Match 25.3%; Score 208.6; DB 21; Length 72750;  
 Best Local Similarity 68.9%; Pred. No. 1.2e-36;  
 Matches 286; Conservative 0; Mismatches 129; Indels 0; Gaps 0;

QY 272 ccaagctgagatgacctcgtcaactggccgcggagagttctctgtggcccatgaccttcg 331  
 Db |||||  
 Db 53866 CCAGCGCGGATACGGTCTGTAACATATATGCGTACCGGTTCGTGTGGCGGGTACTTTCG 53807  
 QY |||||  
 QY 332 gcttgacctgctgcgcgtggagatgatcacatggcagaccocccgtacacatggacc 391  
 Db |||||  
 Db 53806 GCTTGGCCTCTCGCGGTGGAAATGATGCACGCGGTATGGCGGTTACGACCTTGACC 53747  
 QY |||||  
 QY 392 gctttggcgtggtctctccgcgcagcccgccagtcgcagtcgatgatcgtgcccgcga 451  
 Db |||||  
 Db 53746 GTTTCGGTATTATTTCCGTCCGTCGCCCGCTCAGCGCGACCTGATGATTGTGGCGGTA 53687

QY 452 cactcaccacaagaatggcccccagcgttcgcaaggtctacaccagatgcccggagccgc 511  
 Db |||||  
 Db 53686 CGCTCACCAATAAAATGGCGCGCCCTGCGCGAGTGTACGCCAGCTCGCGAGCGGC 53627  
 QY 512 gctacgtggtctccatggggagctgcgccaacggaggaggtactaccactattctact 571  
 Db |||||  
 Db 53626 GCTGGGTATTGTCTATGGGCTCATGTGCCAAGCGCGCGCTATTATCACTATTCTTATT 53567  
 QY 572 cgggtggtgaggggctgcaccgcacgtgcccgtggacatctacatcccaggctgccac 631  
 Db |||||  
 Db 53566 CCGTTGTGCGCGGTGCGGACCGCGTCTGCGGTAGATGTTTATGTGCGGGTGTGTCGC 53507  
 QY 632 ctacggccggaggccctgctctctacggcatcctgcagctgcagaggagatcaagcg 686  
 Db |||||  
 Db 53506 CGACTGCGGAAGCCCTGATTACGGCTGATTACAGCTCCACAAAAATCAAGCG 53452

Search completed: February 12, 2002, 17:22:05  
 Job time: 1620 sec

Result No.	Score	Query %		Length	DB	ID	Description
		Match					
1	715.8	86.9	774	11	BG706832	BG706832	602671984
2	706.8	85.8	835	11	BG479937	BG479937	602527193
3	702.4	85.2	799	10	AL580649	AL580649	AL580649
4	699.8	84.9	879	11	BE794266	BE794266	601591551
5	692.8	84.1	738	11	BG451103	BG451103	602519937
6	692.4	84.0	695	11	BG488658	BG488658	602534588
7	682.4	82.8	732	11	BI116730	BI116730	602868416
8	677.8	82.3	742	11	BG765790	BG765790	602739620
9	661.2	80.2	813	11	BG655478	BG655478	602738979
10	652.6	79.2	936	11	BF968344	BF968344	602369415
11	652.4	79.2	732	11	BE795373	BE795373	601592903
12	651.8	79.1	1002	11	BG395342	BG395342	602457879

```
/clone_lib="NIH_MGC_96"
```

```
/clone_lib="NIH_MGC_96"
```











LOCUS B1116730 733 bp mRNA EST 26-JUN-2001  
 DEFINITION 602868415f1 NIH\_MGC\_7 Homo sapiens cDNA clone IMAGE:501179 5',  
 mRNA sequence.  
 ACCESSION B1116730  
 VERSION B1116730.1 GI:114567631  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 733)  
 AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: cypbs-r@mail.nih.gov  
 Tissue Procurement: DCD/DLP  
 cDNA Library Preparation: Ling Hong/Rubin Laboratory  
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 Cloning Strategy: Lambda ZAP  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Cloning Vector: pCMVcat  
 Clone Distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at: [image.llnl.gov](http://image.llnl.gov)  
 Plate: L1CM1828 row: o column: 04  
 High quality sequence stop: 670.  
 Location/Qualifiers  
 1. 733  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:501179"  
 /clone\_lib="NIH\_MGC\_7"  
 /tissue\_type="small cell carcinoma"  
 /cell\_line="MGC3"  
 /lab\_host="DH10B (phage-resistant)"  
 /note="Organ: lung; Vector: pCMVcat; Site: 1: XhoI; Site: 2:  
 EcoRI; cDNA made by oligo-dT priming. Directionally  
 cloned into EcoRI/XhoI sites using the following 5'  
 adaptor: GGACACGAG(G). Size-selected >500bp for average  
 insert size 1.8kb. Library constructed by Ling Hong in  
 the laboratory of Gerald M. Rubin (University of  
 California, Berkeley) using ZAP-cDNA synthesis kit  
 (Stratagene) and Superscript II RT (Life Technologies)."  
 BASE COUNT 130 a 253 c 230 g 120 t  
 ORIGIN  
 Query Match 82.8%; Score 682.4; DB 11; Length 733;  
 Best Local Similarity 97.6%; Pred. No. 3.9e-122;  
 Matches 703; Conservative 0; Mismatches 16; Indels 1; Gaps  
 QY 105 ggcggcctccggatccttggtcgctccagctggggcggctgtgcaggcacgag 164  
 Db 2 GCGCGCTTCGGATCCTTGCTGTGGCTCCAGCGTGGCGCTGTGCAGGACGAGG 61  
 QY 165 tgtccatcagacgtgtgccacgatgcccagcagcagccacgcctgcctgcaaggc 224  
 Db 62 TGTCATCAGACGTGGGCCACCATGATGCCACAGACACCGCTGCCCTGCCAAAGC 121  
 QY 225 cagagccgtggctcccaaacccagcagcggggcgagtgtggtggccaaagctggatga 284  
 Db 122 CAGAGCGGTGGCTCCAAACCAGCAGCGGGGCGAGTATGTGTGCCAAGCTGATGA 181  
 QY 285 cctcgtcaactggccgcgcggagttctctgtgcccctgacctcgccctggcctgctg 344  
 Db 182 CCGTGTCAACTGGGCCGCGGAGTTCTGTGTGGCCATGACCTTCGGCCTTCGCTGCTG 241  
 QY 345 cgcgtggagatgatgcacatggcagcccccgcctacgacatggaccgcttggcgctgtg 404  
 Db 242 CGCGCTGGAGATGTCACATGCGCAGCACCCCGCTACGACATGGACCGCTTGGCTGTGT 301  
 QY 405 cttccgcgccagccgcgcgcagtcacgtccagtcacgtatcgtgcccggcacatccacaaca 464  
 Db 302 CTTCGCGCCAGCCGCGCAGTCCGACGCTATGATCGTGGCGGCACACTCACCACAA 361





cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)." 127 a 244 c 240 g 125 t

	BASE COUNT	ORIGIN	Query Match Best Local Similarity 79.2%; Score 652.4; DB 11; Length 736; Matches 719; Conservative 0; Mismatches 16; Indels 7; Gaps 5;
QY	73	gccaagatggcggctgtcagtcttcctggccctgcggcggttccggalcotttgctgcgc	132
Dd	2	GGCAAGATGGCGGTCTCTCASCCTCCGG-CTGCGCGGTTCCGGATCTCGGTGTGGCG	60
QY	133	tccaggctgggccggctgtgcaggcacgaggtgtccatcagacgctgccaccgatgpc	192
Dd	61	TCCAGCGTGGCGCTGGCTGTCCAGGACGACGAGGTGTCCATCAGAGCGTGCCACCAGTGGC	120
QY	193	ccaagcagcaccccagctgccttcgcaaaagccagagccgtgctccccaaaaccagcagc	252
Dd	121	CCAAGCAGCACCCAGCCTTCCTGCCAAAGCCAGAGCCGTGGCTCCCCAACCCAGCAGC	180
QY	253	cggggcgagtagtgttgcccaaactggatgacctgtcaactggccgcgcggagttct	312
Dd	181	CGGGGCGAGTAGTGGTGGCCAACTGGATGAACCTGCTCAACTGGCCCGCCGGAGTTCT	240
QY	313	ctgtggcccatgacctcggcctggcctgtggccgtggagatgatgcattgacgaca	372
Dd	241	CTGTGSCCATGACCTTCGGCCTGGCCTGCTGGCGCTGGAGATGATGCACATGCGACGA	300
QY	373	ccocgtaacatggaacgcgtttggcgtggttttcgcgcgaagccgcgcagtcgcac	432
Dd	301	CCOCGCTAGCATGACGACCGCTTTGGCGTGCTTTCGCGCCACGCCCGCCAGTCCGAC	360
QY	433	gtcatgatgtggccgcgcacactcaccacaagaatggccccagcgtctgcgaaggtctac	492
Dd	361	GTCATGATCGTGGCGCGGCACACTCACCACAAGAATGSCCCCACGCGTTGCCAAGGTCTAC	420
QY	493	gaccagatgcggagccgcgtacgttggtctccatggggagctgcgcoaacggagaggc	552
Dd	421	GACCAAGTCCCGAGCCCGCTAGCTGGTCTCCATGGGGAGCTGCGCCAACGAGGAGGC	480
QY	553	tactaccattctactcgttggtgaagggtgcgaccgcactgctgcctggacatc	612
Dd	481	TACTACCATTATTCCTACTCGTGGTGANGGGCTGCACCGCATCTGTCGCCGTGGACATC	540
QY	613	tatatccaggctgccaccctacggccgagggccctgtctacggcatcctgagctgaag	672
Dd	541	TATATCCAGGTGCCACCTACGG-CGAGGCCCTGTCTCTACGGCATCTCTGCAGCTGCCAG	599
QY	673	aggaatatacggagacggaggtcagatctggtaccgaggtgagcgcgcgcgc	732
Dd	600	AGGAATCAAAGGGAGGAGGCTGAGATCTGTACCCGCAAGTAG--CCCGGGCGC	656
QY	733	cgcgcggagcctgtccgcgttcctgtccccagcctgtgttcccgtagagttgcaa	792
Dd	657	GGCGCGCGGAGCTTCGCCGCTCTGTGCCAG-CTGCTTGTGTCCCGTGAGTTGTC-A	714
QY	793	taaaactgcctcgggcaaaaa 814	
Dd	715	TAAACCTGCTCGGAAAAAAA 736	

RESULT 12  
EG395342 BG395342 1002 bp mRNA EST 12-MAR-2001  
LOCUS 602457879F1 NIH\_MGC\_16 Homo sapiens cDNA clone IMAGE:4580538 5',  
DEFINITION mRNA sequence.  
ACCESSION BG395342





found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov  
 Plate: LLCMI771 row: m column: 03  
 High quality sequence stop: 770.

## FEATURES

Location/Qualifiers  
 1. 788  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone\_lib="NIH\_MGC\_19"  
 /tissue\_type="neuroblastoma"  
 /lab\_host="DH10B (phage-resistant)"  
 /note="Organ: brain; Vector: pOTB7; Site\_1: XhoI; Site\_2:  
 EcoRI; cDNA made by oligo-dT priming. Directionally  
 cloned into EcoRI/XhoI sites using the following 5'  
 adaptor: GGACGAG(G). Library constructed by Ling Hong  
 in the laboratory of Gerald M. Rubin (University of  
 California, Berkeley) using ZAP-cDNA synthesis kit  
 (Stratagene) and Superscript II RT (Life Technologies).  
 Note: this is a NIH\_MGC Library."  
 147 a 254 c 254 g 133 t

BASE COUNT  
 ORIGIN

Query Match 78.5%; Score 647.2; DB 11; Length 788;  
 Best Local Similarity 92.7%; Pred. No. 2.3e-115;  
 Matches 714; Conservative 0; Mismatches 48; Indels 8; Gaps 3;

QY 55 gttgtgtgaagcgcaggaagcagatggcgtgtctcagctcctggcctgcgcgcttc 114  
 Db 1 ggtgtgtgaagcgcaggaagc--AAGATGGGGTGTCTGACCTCTGCGCTTCCGCGCTC 58  
 QY 115 cggatccttggctgcctcagcgtggcgcgctgtgcagcagcaggtgtccatcag 174  
 Db 59 CGGATCCTTGCTGCGCTCCAGCGTGGCGCGCTGTGCAGGCACAGAGTGTCCATCAG 118  
 QY 175 agcgtgcccacagatggcccaagcagcaccagcctgcctgcgcaagcagcagcgtg 234  
 Db 119 AGCGTGGCCACCGATGGCCCAAGCAGCACCAGCCCTGCGCTGCAAGGCCAGCGCTG 178  
 QY 235 gctcccaaacccagcagcgcggcgagatgtgtgtggcgaagcttgatgacctcgtcaac 294  
 Db 179 GCTCCCAACCCAGCAGCCGGCGAGTATGTGTGGCCCAAGCTGGATGATCACTCTGAC 238  
 QY 295 tggcgcgcgcggagttctctgtggcccatgaccttgcgcctggcctgtgcgcctggag 354  
 Db 239 TGGCCCGCGCGGAGTTCGTGTGCGCCATGACCTTCGGCTGCGCTGCTGCGCGTGGAG 298  
 QY 355 atgatgcacatggcagcaccgcctacgacatgagcgcgtttggcgtgtgttcgcgcgc 414  
 Db 299 ATGATGCACATGGCAGCACCCTTACGACATGACCGCTTTGGCGTGTCTTCCGCGCC 358  
 QY 415 agccgcgcagtcgcagcgtcatgctgtggcgcgcgcacactcaccacaaagatggcccca 474  
 Db 359 AGCCCGCGCAGTCCGAGCTCATGTCGTGGCGGCACACTCACCACAAAGATGCCCCA 418  
 QY 475 gcgttcgcaggtctacgaccagatgcgcgagcgcgtacgtggttcctccatggggagc 534  
 Db 419 CGGCTTCGCAAGGCTACGACCAAGATGCCGAGCCCGCTACGTGTGTCTCCATGGGGAGC 478  
 QY 535 tgcgcacaagcagagaggtactaccactattcctactcgtgtgtgagggcgtgcgaccgc 594  
 Db 479 TGGCCCAACCGAGGAGGCTACTACCACTATTCTACTCGTGTGTGAGGGGCTGCCACGCG 538  
 QY 595 atcgtgcgcgtgacatctacatccaggctgcgccacctacgcgcgagggccctgtctac 654  
 Db 539 ATCGTGCCCGTGGACATCTACATCCCAAGGTGCC--ACCTAGCGGCGAGGCCCTGCTTAC 596  
 QY 655 ggcacctgcagctgcagagaagatcaagcggagcggagcagagctgcagatcgtgtaccgc 714  
 Db 597 GGCATCTGCTGAGTGCAGAGAGAGATCAAGCGGGAGCGGAGGCTGCAGATCTGGTACCGC 656  
 QY 715 aggtagc 774

Db 657 AGGTAGGCCCG---CGCCCGCGCGCGCGAGCTGTGCCCTGTGTCCCGAGCTGTGTG 712  
 QY 775 gtcccgtaggtgtcacaataaacctgcctcggcgcaaaaaaataaaaaa 824  
 Db 713 TCCCGGTGAGTTGTCAATAAACCTGCTCGGAACACAAAAAATAAAA 762

## RESULT 15

LOCUS BG419411 910 bp mRNA EST 14-MAR-2001  
 DEFINITION 602446102F1 NIH\_MGC\_14 Homo sapiens cDNA clone IMAGE:4584528 5',  
 mRNA sequence.  
 ACCESSION BG419411  
 VERSION BG419411.1 GI:13325917  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 910)  
 AUTHORS NIH-MGC http://mgc.nci.nih.gov/.  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: cyapbs-remail.nih.gov  
 Tissue Procurement: DCTD/DTP  
 CDNA Library Preparation: Ling Hong/Rubin Laboratory  
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
 http://image.llnl.gov  
 Plate: LLCMI310 row: d column: 01  
 High quality sequence stop: 702.

FEATURES  
 source

Location/Qualifiers  
 1. 910  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone\_lib="NIH\_MGC\_14"  
 /tissue\_type="renal cell adenocarcinoma"  
 /lab\_host="DH10B (phage-resistant)"  
 /note="Organ: kidney; Vector: pOTB7; Site\_1: XhoI; Site\_2:  
 EcoRI; cDNA made by oligo-dT priming. Directionally  
 cloned into EcoRI/XhoI sites using the following 5'  
 adaptor: GGACGAG(G). Size-selected >500bp for average  
 insert size 1.8kb. Library constructed by Ling Hong in  
 the laboratory of Gerald M. Rubin (University of  
 California, Berkeley) using ZAP-cDNA synthesis kit  
 (Stratagene) and Superscript II RT (Life Technologies)."  
 BASE COUNT 195 a 279 c 278 g 156 t 2 others  
 ORIGIN

Query Match 78.3%; Score 645.6; DB 11; Length 910;  
 Best Local Similarity 94.7%; Pred. No. 4.6e-115;  
 Matches 713; Conservative 0; Mismatches 30; Indels 10; Gaps 4;

QY 72 ggccaagatggcgtgctcagctcctgcgcgcgcgttcgcggtcctggtctgcg 131  
 Db 1 GGCCAGATGGCGGTGCTGTCAGCTCTCTGGCTGCGCGCTTCCGGATCCTTGGTCTCGG 60  
 QY 132 ctccagcgtggcccgctgtgcagcagcaggtgtccatcagagcgtggccaccgattg 191  
 Db 61 CTCAGCGTGGCGCTGGCTGTGCAGGCAGAGGTGTCCATCAGAGCGTGGCCACCGATGG 120  
 QY 192 ccaaacagcaccagcctgcctgcgaagccagcgcgtggtcccaaccagcag 251  
 Db 121 CCAGAGCAGCACCAGCGCTGCGCTGCCAAGGCCAGAGCGGTGGCTCCCAACCAGCAG 180  
 QY 252 ccggggcgagtatgtgtggccaaagctggatgacctcgtcaactggcgccgcgagttc 311  
 Db 181 CCGGGCGAGTATGTGGTGCCCAAGCTGGATGACCTCGTCAACTGGGCCCGCGGAGTTC 240

```
QY 312 tctgtgcccacatgaaccttcggccttgccctgtgcccgtgagatgatgcacatggcagc 371
Db 241 TCCTGTGGCCCATGACCTTCGGCTGGCTGCTGCGCTGGAGATGATGCACATGGCAGC 300
QY 372 accccgctacgacatggaccgtttgggtgtgttcctccgcccagcccgccagtcgca 431
Db 301 ACCCCGCTACGACATGGACCGTTTGGGTGTGTCTCCGCGCCAGCCGCGCCAGTCCGA 360
QY 432 cgtcatgatcgtggccggcacactcaccacaagaatggccccagcgttcgcgaaggtcta 491
Db 361 CGTCATGATCGTGGCGGCGCACACTACCAACAAGATGCCCCAGCGCTTCGCAAGGTCTA 420
QY 492 cgaccagatgcggagagccgcgtacgtgtgttcctccatggggagcctgcgccaacgagaggg 551
Db 421 CGACCAGATGCCGGAGCCGCGTACGTGTCTCCATGGGGAGCTGCCGCCAACGGAGGAGG 480
QY 552 ctactaccactattcctactcgtgtgtgtgaggggtgcgaccgcacgtgcccgtggacat 611
Db 481 CTACTACCACATATTCCTACTCGTGTGTGAGGGGCTGCGACCGCATCGTGCCCGTGACAT 540
QY 612 ctacatcccaggctgcccacctacgcccagggccctgtctctacggcatcctgcagctgca 671
Db 541 CTACATCCCAGGCTG--CCACCTACGCG--GAGGCGCTGCTTACGGCATCCTGCGCTGCA 597
QY 672 gaggaagatcaagcggagcggagcgtgcagatctggtaccgcaggtagcgcgcgcgcg 731
Db 598 GAGGAAGATCAAGCGGGAGCGAGGCTGCAGATCTGGTA--CGCAGGTAG-----CGCGC 650
QY 732 ccgcgcggagcctgtgcgcctcctgtcccccagcctgtgttcccgtgaggttgc 791
Db 651 GCGCGCGGAGGCTGTGCGCCCTCCTGTCCNCCAGCTGTTGTTCGCCGTGAGGTGTC 710
QY 792 ataaacctgccctcgggcaaaaaaaaaaaaaa 824
Db 711 ATRACCTGGCTCGCGGAAAAAAAAAAAAAAAAA 743
```

Search completed: February 12, 2002, 17:41:38  
Job time: 2873 sec

Result No.	Score	Query Match	Length	DB	ID	Description
1	1118	100.0	213	4	Q9BV17	Q9bv17 homo sapien
2	936.5	83.8	224	11	Q9DC70	Q9dc70 mus musculu
3	815.5	72.9	214	4	Q9HK35	Q9hk35 homo sapien
4	753.5	67.4	221	5	Q9VXK7	Q9vxk7 drosophila
5	743	66.5	212	5	Q9VAK5	Q9vak5 drosophila
6	731	65.4	213	10	Q9SP38	Q9sp38 lupinus lut
7	721.5	64.5	192	10	Q9LKG9	Q9lkg9 lupinus lut
8	716	64.0	158	8	Q9G8U4	Q9g8u4 rhodomonas
9	712.5	63.7	210	10	Q9LKH4	Q9lkh4 lupinus lut
10	702	62.8	210	3	Q9UUT7	Q9uut7 yarrowia li
11	699	62.5	158	8	Q9TCA4	Q9tca4 nephroselmi
12	654	58.5	191	2	Q9PGJ4	Q9pgj4 xylella fas
13	642.5	57.5	193	2	Q9A6X1	Q9a6x1 caulobacter
14	613	54.8	160	2	Q9K1C2	Q9k1c2 neisseria m
15	605	54.1	160	2	Q9JX79	Q9jx79 neisseria m
16	598	53.5	162	8	Q9XMU2	Q9xmu2 tetrahymena
17	497.5	44.5	159	2	Q9S851	Q9s851 helicobacte
18	496.5	44.4	159	2	Q9ZJW6	Q9zjw6 helicobacte
19	474	42.4	181	2	Q9RU87	Q9ru87 deिनococcus

DE	MY017 PROTEIN.
OS	Homo sapiens (Human).
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NCBI_TaxID=9606;	
[1]	
RN	
RP	SEQUENCE FROM N.A.
RC	TISSUE=BRAIN;
RL	Mao Y.M., Xie Y., Huang X.Y., Ying K., Dai J.L.;
RL	Submitted (APR-1998) to the EMBL/GenBank/DBJ databases.
DR	EMBL; AF060512; AAG43130.1;
DR	InterPro; IPR002096; Complex_20kD.
DR	Pfam; PF01058; oxidored_46; 1..20Kd.
SQ	SEQUENCE 214 AA; 23622 MW; 8DCE7E13F2FC7B3 CRC64;

  

Query Match	72.9%; Score 815.5; DB 4; Length 214;
Best Local Similarity	75.3%; Pred. No. 3.7e-68;
Matches 171; Conservative	7; Mismatches 22; Indels 27; Gaps:

  

QY	1	MAVLSAPGLRGFRILGILRSSVGPAVGARGVHQSVAITDGPSSTQPALPKARAVAPKPSRSRG 60
DB	1	MAVLSAPGLRGFRILGILRSSVGLAVGARGVHQSVAITDGPSSTQPALPKARAVAPKPSRSRG 60
QY	61	EYVAKLDLNVNARRSLPMTFGLCACAVEMHMAAPRYDMDFGVVFRAASPROSDVM 120
DB	61	EYVAKLDLNVNARRSLPMTFGLCACAVEMHMAAPRYDMDFGVVFRAASPROSDVM 120
QY	121	IVAGILTNNKMAPALKVVDOMPFRYYVSGSANGGGYYHYSVVRGCDRIVPVDI-- 178
DB	121	IVAGILTNNKMAPALK---SPRRP-----DAGAALRGLHGELQRGRULPLFLLG 167
QY	179	-----YIPGCPPTAEALLGILOLRKKIKRRRLQIYWR 213
DB	168	GEGLPHPRHARGHLPRPTTAEALLGILOLRKKIKRRRLQIYWR 214

  

RESULT	4
ID	Q9VXK7 PRELIMINARY; PRT; 221 AA.
QC	Q9VXK7
AC	01-MAY-2000 (TREMBLrel. 13, Created)
DT	01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DE	01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE	CG9172 PROTEIN.
GN	CG9172.
OS	Drosophila melanogaster (Fruit fly).
OC	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC	Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC	Ephydroidea; Drosophilidae; Drosophila.
ON	NCBI_TaxID=7227;
RN	[1]
RP	SEQUENCE FROM N.A.
RC	STRAIN=BERKELEY;
RX	MEDLINE=20196006; PubMed=107311132;
RA	Adams M.D., Celiker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA	Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
RA	George R.A., Lewis S.B., Richards S., Ashburner M., Henderson S.N.,
RA	Sutton G.G., Wortman J.R., Vandell M.D., Zhang Q., Chen L.X.,
RA	Branan R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA	Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA	Abriel J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA	Balaw R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA	Beeson K.Y., Benos P.V., Bereman J.P., Bhattacharya S., Bolshakov S.,
RA	Borkova D., Botchan M.R., Bouck J., Brockstein P., Brotter P.,
RA	Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA	Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA	de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA	Dodson K., Dou P.L.E., Downes M., Dugan-Rocha S., Dunkow B.C., Dunn P.,
RA	Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA	Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA	Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA	Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,



QY	148	VSMGSCANGGGYHYHSYSVRGCDRIVPVDIYIPGCPPTAEALLGILQKRIKRRRL	207
Db	127	ISMGSCANGGGYHYHSYSVRGCDRIVPVDIYVPGCPPTAEALLGILQKRIKRRKF	186
QY	208	QIWYRR 213	
Db	187	LLWWTk 192	
RESULT	8		
ID	Q9G8U4	PRELIMINARY;	PRT; 158 AA.
AC	Q9G8U4;		
DT	01-MAR-2001 (TReMBLrel. 16, Created)		
DT	01-MAR-2001 (TReMBLrel. 16, Last sequence update)		
DT	01-JUN-2001 (TReMBLrel. 17, Last annotation update)		
DE	NADH DEHYDROGENASE SUBUNIT 10 (EC 1.6.5.3).		
GN	NAD10.		
OS	Rhodomonas salina.		
OC	Mitochondrion.		
OC	Eukaryota; Cryptophyta; Cryptomonadaceae; Rhodomonas.		
OX	NCBI_TaxID=52570;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RA	Burger G., Lang B.F., Maier U.G., McFadden G.I., Gray M.W.;		
RT	"Algae with secondary chloroplasts have mitochondria that originate from the host."		
RL	Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.		
DR	EMBL; AF288090; AGI17753.1;		
DR	InterPro; IPR002096; Complex1_20kD.		
DR	Pfam; PF01058; oxidored_46; 1.		
DR	PROSITE; PS01150; COMPLEX1_20K; 1.		
KW	Oxidoreductase; Mitochondrion.		
SQ	SEQUENCE 158 AA; 17819 MW; 0798BB29D71611C4 CRC64;		
Query Match	64.08;	Score 716;	DB 8; Length 158;
Best Local Similarity	79.99;	Pred. No. 4.6e-59;	
Matches 123;	Conservative 22;	Mismatches 9;	Indels 0; Gaps
QY	58	SRGEYVAKLDLDLVNWARSSLPMTFGLACCAVENMHMAAPRYDMDRFGVVFPRASPRQS	117
Db	3	NKTKFVSKIDLLNWARRGSLPMTFGLACCAVENMHAGASRYDFDRFGIIFRASPQS	62
QY	118	DWMIYAGILTINKMAPALKRYQDMPERYVSWGSCANGGGYHYHSYSVRGCDRIVPVD	177
Db	63	DIMIVAGILTINKMAPALKRYQDMPERYVSWGSCANGGGYHYHSYSVRGCDKIVPVD	122
QY	178	IYIPGCPPTAEALLGILQKRIKRRRLQWY 211	
Db	123	VFVPGCPPTAEALLGILQKRIKRTVQTWF 156	
RESULT	9		
ID	Q9LKH4	PRELIMINARY;	PRT; 210 AA.
AC	Q9LKH4;		
DT	01-OCT-2000 (TReMBLrel. 15, Created)		
DT	01-OCT-2000 (TReMBLrel. 15, Last sequence update)		
DT	01-JUN-2001 (TReMBLrel. 17, Last annotation update)		
DE	NADH-UBIQUINONE OXIDOREDUCTASE SUBUNIT PSST.		
OS	Lupinus luteus (Yellow lupine).		
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;		
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;		
OC	eurosid 1; Fabales; Fabaceae; Papilionoideae; Lupinus.		
OX	NCBI_TaxID=3873;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RA	Piekna D., Sikorski M., Augustyniak H.;		
RT	"Lupinus luteus PSST gene."		
RL	Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.		
DR	EMBL; AF279902; AAF91429.1;		

DR InterPro: IPR002096; Complex1\_20kD.  
DR Pfam: PF01058; oxidored\_46; 1.  
DR PROSITE; PS01150; COMPLEX1\_20K; 1.  
KW Ubiquinone.  
SQ SEQUENCE 210 AA; 23455 MW; C5E1AE2F48C412E3 CRC64;

Query Match 63.7%; Score 712.5; DB 10; Length 210;  
Best Local Similarity 69.9%; Pred. No. 1.4e-58;  
Matches 130; Conservative 21; Mismatches 22; Indels 13; Gaps 2;

Qy 41 STQPAL-----PKARAVAPKPS-----RGEYVYVAKLDDLNVNRRSSLPWMTFGLA 87  
Db 25 TLESLSPEATTPYSPRPSPSSPAGISKTAEFVISKVDDLNNWARRGSWPMTFGLA 84  
Qy 88 CCAYEMHMAAPRYDMDFGVFRASPSQSDVMIVAGTITNKNAPALRVYDQMPPEPRV 147  
Db 85 CCAYEMHMTGAARYDLDLFRGFIIFRSPQSDCMIVAGTITNKNAPALRVYDQMPPEPRV 144  
Qy 148 VSMGSCANGGYHYHSYVVGCDRIVPVDIYIPGCPPTAEALLYGILQLOKRIKREL 207  
Db 145 ISMGSCANGGYHYHSYVVGCDRIVPVDIYIPGCPPTAEALLYGILQLOKRIKREL 204  
Qy 208 QIYRR 213  
Db 205 LLWTK 210

RESULT 10  
Q9UUT7 PRELIMINARY; PRT; 210 AA.  
AC Q9UUT7;  
DT 01-MAY-2000 (TEMBLrel. 13, Created)  
DT 01-MAY-2000 (TEMBLrel. 13, Last sequence update)  
DT 01-JUN-2001 (TEMBLrel. 17, Last annotation update)  
DE SUBUNIT NUKM OF PROTEIN NADH:UBIQUINONE OXIDOREDUCTASE (COMPLEX I)  
GN PRECURSOR (EC 1.6.99.3).  
OS Yarrowia lipolytica (Candida lipolytica).  
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
OC Saccharomycetales; Dipodascaceae; Yarrowia.  
OX NCBI\_TaxID=4952;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=E 150;  
RA Djafarzadeh Andabali R., Kerscher S., Zwicker K., Radermacher M.,  
RA Lindahl M., Schaeffer H., Brandt U.;  
RT "Biophysical and Structural Characterization of Proton-Translocating  
NADH-Dehydrogenase from Yarrowia lipolytica";  
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AJ250340; CAB65525.1; -  
DR InterPro: IPR002096; Complex1\_20kD.  
DR Pfam: PF01058; oxidored\_46; 1.  
DR PROSITE; PS01150; COMPLEX1\_20K; 1.  
KW Signal; Oxidoreductase; NAD.  
FT SIGNAL 1 27  
FT CHAIN 28 210  
FT SUBUNIT NUKM OF PROTEIN NADH:UBIQUINONE  
FT OXIDOREDUCTASE (COMPLEX I).  
SQ SEQUENCE 210 AA; 23430 MW; C976335ECFD686ED CRC64;

Query Match 62.8%; Score 702; DB 3; Length 210;  
Best Local Similarity 63.8%; Pred. No. 1.3e-57;  
Matches 125; Conservative 29; Mismatches 32; Indels 10; Gaps 1;

Qy 18 RSSVGPVQARGVHQSVATDGPSTPALPKARAVAPKPSRGEYVYVAKLDDLNVNRRS 77  
Db 25 RAYSAPAGTPRVSSSMETDFPLPSQ-----QKPSAVDYTLTLDVANWARG 74  
Qy 78 SLWPTFTGLACCAVEMHMAAPRYDMDFGVFRASPSQSDVMIVAGTITNKNAPALRV 137  
Db 75 SFWPTFTGLACCAVEMHMAAPRYDMDFGVFRASPSQSDVMIVAGTITNKNAPALRV 134

Qy 138 YDQMPPEPRYVSMGSCANGGYHYHSYVVGCDRIVPVDIYIPGCPPTAEALLKGLQ 197  
Db 135 YDQMPPEPRYVSMGSCANGGYHYHSYVVGCDRIVPVDIYIPGCPPTAEALLKGLQ 194  
Qy 198 QRTKRRERLQIYRR 213  
Db 195 QRTKRRERLQIYRR 210

RESULT 11  
Q9TCA4 PRELIMINARY; PRT; 158 AA.  
AC Q9TCA4;  
DT 01-MAY-2000 (TEMBLrel. 13, Created)  
DT 01-MAY-2000 (TEMBLrel. 13, Last sequence update)  
DT 01-JUN-2001 (TEMBLrel. 17, Last annotation update)  
DE NADH DEHYDROGENASE SUBUNIT 10 (EC 1.6.5.3).  
GN NAD10.  
OS Nephroselmis olivacea.  
OG Mitochondrion.  
OC Eukaryota; Viridiplantae; Chlorophyta; Prasinophyceae;  
OC Chlorodendales; Chlorodendraceae; Nephroselmis.  
OX NCBI\_TaxID=31312;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=NIES-484;  
RA Turmel M., Lemieux C., Burger G., Lang B.F., Otis C., Plante I.,  
RA Gray M.W.;  
RT "The complete mitochondrial DNA sequences of Nephroselmis olivacea and  
RT Pedinomonas minor: two radically different evolutionary patterns  
RL within the green algae";  
RN Plant Cell 11:1717-1729(1999).  
RP SEQUENCE FROM N.A.  
RC STRAIN=NIES-484;  
RA Turmel M., Lemieux C., Burger G., Lang B.F., Otis C., Plante I.,  
RA Gray M.W.;  
RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF110138; AAF03193.1; -  
DR InterPro: IPR002096; Complex1\_20kD.  
DR Pfam: PF01058; oxidored\_46; 1.  
DR PROSITE; PS01150; COMPLEX1\_20K; 1.  
KW Oxidoreductase; Mitochondrion.  
SQ SEQUENCE 158 AA; 17781 MW; CE82530D4724EB3F CRC64;

Query Match 62.5%; Score 699; DB 8; Length 158;  
Best Local Similarity 78.2%; Pred. No. 1.8e-57;  
Matches 122; Conservative 19; Mismatches 15; Indels 0; Gaps 0;

Qy 58 SRGEYVYVAKLDDLNVNRRSSLPWMTFGLACCAVEMHMAAPRYDMDFGVFRASPROS 117  
Db 3 NNAEFLTGIDSLNNWARRKSGWMTFGLACCAVEMHMAAPRYDMDFGVFRASPROS 62  
Qy 118 DVMIVAGTITNKNAPALRVYDQMPPEPRYVSMGSCANGGYHYHSYVVGCDRIVPVD 177  
Db 63 DVMIVAGTITNKNAPALRVYDQMPPEPRYVSMGSCANGGYHYHSYVVGCDRIVPVD 122  
Qy 178 IYIPGCPPTAEALLYGILQLOKRIKRRERLQIYRR 213  
Db 123 IYIPGCPPTAEALLYGILQLOKRIKRRERLQIYRR 158

RESULT 12  
Q9PGJ4 PRELIMINARY; PRT; 191 AA.  
AC Q9PGJ4;  
DT 01-OCT-2000 (TEMBLrel. 15, Created)  
DT 01-OCT-2000 (TEMBLrel. 15, Last sequence update)  
DT 01-JUN-2001 (TEMBLrel. 17, Last annotation update)  
DE NADH-UBIQUINONE OXIDOREDUCTASE, N006 SUBUNIT.  
GN XF0306.

OS Xylella fastidiosa.  
 OC Bacteria; Proteobacteria; gamma subdivision; Xanthomonas group;  
 OC Xylella.  
 OX NCBI\_TaxID=2371;  
 RN [1]  
 RN SEQUENCE FROM N.A.  
 RC STRAIN=9A5C;  
 RX MEDLINE=20365717; PubMed=10910347;  
 RA Simpson A.J.G., Reinach F.C., Artuda P., Abreu F.A., Acencio M.,  
 RA Alvarenga R., Alves L.M.C., Araya J.E., Baia G.S., Baptista C.S.,  
 RA Barros M.H., Bonaccorsi E.D., Bordin S., Bove J.M., Briones M.R.S.,  
 RA Bueno M.R.P., Canaro A.A., Camargo L.E.A., Carraro D.M., Carrier H.,  
 RA Colauto N.B., Colombo C., Costa F.F., Costa M.C.R., Costa-Neto C.M.,  
 RA Coutinho L.L., Cristofani M., Dias-Neto E., Docena C., El-Borri H.,  
 RA Facincani A.P., Ferreira A.J.S., Ferreira V.C.A., Ferro J.A.,  
 RA Fraga J.S., Franca S.C., Franco M.C., Frohme M., Furlan L.R.,  
 RA Garnier M., Goldman G.H., Goldman M.H.S., Gomes S.L., Gruber A.,  
 RA Ho P.L., Hohenisel J.D., Junqueira M.L., Kemper E.L., Kitajima J.P.,  
 RA Krieger J.E., Kuramae E.E., Laigret F., Lambais M.R., Leite L.C.C.,  
 RA Lemos E.G.M., Lemos M.V.F., Lopes S.A., Lopes C.R., Machado J.A.,  
 RA Machado M.A., Madeira A.M.B.N., Madeira H.M.F., Marino C.L.,  
 RA Marques M.V., Martins E.A.L., Martins E.M.F., Matsukuma A.Y.,  
 RA Menck C.F.M., Miracca E.C., Miyaki C.Y., Monteiro-Vitorello C.B.,  
 RA Moon D.H., Nagai M.A., Nascimento A.L.T.O., Netto L.E.S.,  
 RA Nhani A., Jr., Nobrega F.G., Nunes L.R., Oliveira M.A.,  
 RA de Oliveira M.C., de Oliveira R.C., Palmieri D.A., Paris A.,  
 RA Peixoto B.R., Pereira G.A.G., Pereira H.A., Jr., Pesquero J.B.,  
 RA Quaggio R.B., Roberto P.G., Rodrigues V., de Rosa A.J.M.,  
 RA de Rosa V.E., Jr., de Sa R.G., Santelli R.V., Sawasaki H.E.,  
 RA da Silva A.C.P., da Silva A.M., da Silva F.R., Silva W.A., Jr.,  
 RA da Silva J.F., Silvestri M.L.E., Siqueira W.J., de Souza A.A.,  
 RA de Souza A.P., Terenzi M.F., Truffi D., Teai S.M., Tshako M.H.,  
 RA Vallada H., Van Sluys M.A., Verjovski-Almeida S., Vettore A.L.,  
 RA Zago M.A., Zatz M., Meidanis J., Setubal J.C.;  
 RT "The genome sequence of the plant pathogen Xylella fastidiosa";  
 RL Nature 406:151-159(2000).  
 DR EMBL: AE003884; AAF8317.1;  
 DR InterPro; IPR002096; Complexl\_20KD.  
 DR Pfam; PF01058; oxidored\_q6; 1.  
 DR PROSITE; PS01150; COMPLEXI\_20K; 1.  
 KW Complete proteome.  
 SQ SEQUENCE 191 AA; 20858 MW; 7A14CC80A8857CF5 CRC64;

Query Match 58.5%; Score 654; DB 2; Length 191;  
 Best Local Similarity 72.1%; Pred. No. 3.4e-53;  
 Matches 119; Conservative 14; Mismatches 20; Indels 12; Gaps 1;  
 QY 39 PSSTOPALPKARAVAPKPSRGEYVYVAKLDLVNKAARSLWPMTFLGACCAVEMHMAA 98  
 Db 33 POGESPLQKG-----VYTSVDALLNWTGSMWPTFLGACCAVEMHMAA 80  
 QY 99 PRYDMDFGVFRASPRQSDVMIVAGTLTNKMAPALRKVDQMPEPRYVSMGSCANGGG 158  
 Db 81 ARLDLDYGVFRPSRQSDVMIVAGTLVNKMAPALRKVDQMPPDKWVSMGSCANGGG 140  
 QY 159 YVHYSVYVRCGRVVDVYIPGCPPTAEALYGILOLQKKIR 203  
 Db 141 YVHYSVYVRCGRVVDVYIPGCPPTAEALYGILOLQKKIR 185  
 RESULT 13  
 ID Q9A6X1 PRELIMINARY; PRT; 193 AA.  
 AC Q9A6X1;  
 DT 01-JUN-2001 (TReMBLrel. 17, Created)  
 DT 01-JUN-2001 (TReMBLrel. 17, Last sequence update)  
 DT 01-JUN-2001 (TReMBLrel. 17, Last annotation update)  
 DE NADH DEHYDROGENASE I, B SUBUNIT.  
 GN CC1955.  
 OS Caulobacter crescentus.  
 OC Bacteria; Proteobacteria; alpha subdivision; Caulobacter group;  
 OC Caulobacter.

OX NCBI\_TaxID=69394;  
 RN [1]  
 RN SEQUENCE FROM N.A.  
 RX MEDLINE=21173698; PubMed=11259647;  
 RA Nierman W.C., Feldblyum T.V., Laub M.T., Paulsen I.T., Nelson K.E.,  
 RA Eisen J., Heidelberg J.F., Alley M.R.K., Ohta N., Maddock J.R.,  
 RA Potocka I., Nelson W.C., Newton A., Stephens C., Phadke N.D., Ely B.,  
 RA DeBoy R.T., Dodson R.J., Durkin A.S., Gwinn M.L., Hart D.H.,  
 RA Kolonay J.F., Smit J., Craven M.B., Khouri H., Shetty J., Berry K.,  
 RA Utterback T., Tran K., Wolf A., Vamathevan J., Ermolaeva M., White O.,  
 RA Salzberg S.L., Venter J.C., Shapiro L., Fraser C.M.;  
 RT "Complete genome sequence of Caulobacter crescentus";  
 RL Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141(2001).  
 DR EMBL: AE005868; AAK23930.1;  
 DR TIGR; CC1955;  
 KW Complete proteome.  
 SQ SEQUENCE 193 AA; 21082 MW; C5F57AE6B341862 CRC64;  
 Query Match 57.5%; Score 642.5; DB 2; Length 193;  
 Best Local Similarity 56.1%; Pred. No. 4e-52;  
 Matches 123; Conservative 20; Mismatches 26; Indels 17; Gaps 2;  
 QY 34 VAIDGPSSTQALPKARA-----VAPKPSRGEYVYVAKLDLVNWAARS 77  
 Db 3 VIVGNSSPVPSALSAGRTVEGYDPKLDHPDFDGVSOQLADKG-FITAAADDLITWARTG 61  
 QY 78 SLWPTFTGLACCAVEMHMAAPRYDMDFGVFRASPRQSDVMIVAGTLTNKMAPALRKV 137  
 Db 62 SLWPTFTGLACCAVEMHMAAPRYDMDFGVFRASPRQSDVMIVAGTLTNKMAPALRKV 121  
 QY 138 YQDMPEPRYVSMGSCANGGGYHYYSVVRGCDRVIPVDIYIPGCPPTAEALYGILOL 197  
 Db 122 YQDMPEPRYVSMGSCANGGGYHYYSVVRGCDRVIPVDIYIPGCPPTAEALYGILOL 181  
 QY 198 QKKIR 203  
 Db 182 QKKIR 187  
 RESULT 14  
 ID Q9K1C2 PRELIMINARY; PRT; 160 AA.  
 AC Q9K1C2;  
 DT 01-OCT-2000 (TReMBLrel. 15, Created)  
 DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)  
 DT 01-JUN-2001 (TReMBLrel. 17, Last annotation update)  
 DE NADH DEHYDROGENASE I, B SUBUNIT.  
 GN NMG0242.  
 OS Neisseria meningitidis (serogroup B).  
 OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.  
 OX NCBI\_TaxID=491;  
 RN [1]  
 RN SEQUENCE FROM N.A.  
 RC STRAIN=MC58 / SEROGROUP B;  
 RX MEDLINE=20175755; PubMed=10710307;  
 RA Tettelin H., Saunders N.J., Heidelberg J., Jeffries A.C., Nelson K.E.,  
 RA Eisen J.A., Ketchum K.A., Hood D.W., Peden J.F., Dodson R.J.,  
 RA Nelson W.C., Gwinn M.L., DeBoy R., Peterson J.D., Hickey E.K.,  
 RA Haft D.H., Salzberg S.L., White O., Fleischmann R.D., Dougherty B.A.,  
 RA Mason T., Ciecko A., Parksey D.S., Blair E., Citterone H., Clark E.B.,  
 RA Cotton M.D., Utterback T.R., Khouri H., Qin H., Vamathevan J.,  
 RA Gill J., Scarlato V., Maignani V., Pizzi M., Grandi G., Sun L.,  
 RA Smith H.O., Fraser C.M., Moxon E.R., Rappuoli R., Venter J.C.;  
 RT "Complete genome sequence of Neisseria meningitidis serogroup B strain MC58";  
 RL Science 287:1809-1815(2000).  
 DR EMBL: AE002381; AAP40696.1;  
 DR TIGR; NMB0242;  
 DR InterPro; IPR002096; Complexl\_20kD.  
 DR Pfam; PF01058; oxidored\_q6; 1.  
 DR PROSITE; PS01150; COMPLEXI\_20K; 1.  
 KW Complete proteome.



```
SQ SEQUENCE 160 AA; 17629 MW; F1616423A8D799AF CRC64;

Query Match          54.1%; Score 613; DB 2; Length 160;
Best Local Similarity 73.2%; Pred. No. 9.9e-49;
Matches 104; Conservative 22; Mismatches 16; Indels 0; Gaps 0;

QY 62 YVAKLDLVNWKARRSSLPWMTFGLACCAVEMMHMAAPRYDMDRFGVYVFRASPRQSDVMI 121
Db 11 FITTSADTVLNMRTGSLWPVTFGLACCAVEMMHMAAPRYDMDRFGVYVFRASPRQADLMI 70
QY 122 VAGTITNKMAPALRKRVYDQMPPEPVYVSMGSCANGGGYHYHSYVVRGCDRIYVPDIYIP 181
Db 71 VAGTITNKMAPALRRVYDQLAEPWRVLSMGSCANGGGYHYHSYVVRGADRVVPDVYVP 130
QY 182 GCPPTAEALYGLIQLQKIKR 203
Db 131 GCPPTAEALYGLIQLQKIKR 152

RESULT 15
Q9JX79
ID Q9JX79 PRELIMINARY; PRT; 160 AA.
AC Q9JX79;
DT 01-OCT-2000 (TremBrel. 15, Created)
DT 01-OCT-2000 (TremBrel. 15, Last sequence update)
DT 01-JUN-2001 (TremBrel. 17, Last annotation update)
DE NADH DEHYDROGENASE I CHAIN B (EC 1.6.5.3).
GN NUOB OR NMA0018.
OS Neisseria meningitidis (serogroup A).
OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.
OX NCBI_TaxID=65569;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=22491 / SEROGROUP A / SEROTYPE 4A;
RX MEDLINE=20222556; PubMed=10761919;
RA Parkhill J., Achtman M., James K.D., Bentley S.D., Churcher C.,
RA Klee S.R., Morelli G., Basham D., Brown D., Chillingworth T.,
RA Davies R.M., Davis P., Devlin K., Feltwell T., Hamlin N., Holroyd S.,
RA Jagels K., Leather S., Moule S., Mungall K., Quail M.A.,
RA Rajandream M.A., Rutherford K.M., Simmonds M., Skelton J.,
RA Whitehead S., Spratt B.G., Barrall B.G.;
RT "Complete DNA sequence of a serogroup A strain of Neisseria
RT meningitidis 22491."
RL Nature 404:502-506(2000).
DR EMBL; AL162752; CAB83338.1;
DR InterPro; IPR002096; Complex1_20KD.
DR Pfam; PF01058; oxidored_66; 1.
DR PROSITE; PS01150; COMPLEX1_20K; 1.
KW Oxidoreductase; Complete proteome.
SQ SEQUENCE 160 AA; 17619 MW; 4E9CF64C554799AC CRC64;

Query Match          54.1%; Score 605; DB 2; Length 160;
Best Local Similarity 73.2%; Pred. No. 9.9e-49;
Matches 104; Conservative 22; Mismatches 16; Indels 0; Gaps 0;

QY 62 YVAKLDLVNWKARRSSLPWMTFGLACCAVEMMHMAAPRYDMDRFGVYVFRASPRQSDVMI 121
Db 11 FITTSADTVLNMRTGSLWPVTFGLACCAVEMMHMAAPRYDMDRFGVYVFRASPRQADLMI 70
QY 122 VAGTITNKMAPALRKRVYDQMPPEPVYVSMGSCANGGGYHYHSYVVRGCDRIYVPDIYIP 181
Db 71 VAGTITNKMAPALRRVYDQLAEPWRVLSMGSCANGGGYHYHSYVVRGADRVVPDVYVP 130
QY 182 GCPPTAEALYGLIQLQKIKR 203
Db 131 GCPPTAEALYGLIQLQKIKR 152
```



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

# OM protein - protein search, using sw model

Run on: February 12, 2002, 19:56:10 ; Search time 37.72 Seconds  
(without alignments)  
418.282 Million cell updates/sec

Title: US-09-525-867-1

Perfect score: 1118

Sequence: 1 MAVLSAPGLRGPRILGLRSS.....ILQLQKTKRRRLQIYYRR 213

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 522463 seqs, 74073290 residues

Total number of hits satisfying chosen parameters: 522463

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : 1: /SID52/gcgdata/geneseq/geneseq/AA1980.DAT.\*  
2: /SID52/gcgdata/geneseq/geneseq/AA1981.DAT.\*  
3: /SID52/gcgdata/geneseq/geneseq/AA1982.DAT.\*  
4: /SID52/gcgdata/geneseq/geneseq/AA1983.DAT.\*  
5: /SID52/gcgdata/geneseq/geneseq/AA1984.DAT.\*  
6: /SID52/gcgdata/geneseq/geneseq/AA1985.DAT.\*  
7: /SID52/gcgdata/geneseq/geneseq/AA1986.DAT.\*  
8: /SID52/gcgdata/geneseq/geneseq/AA1987.DAT.\*  
9: /SID52/gcgdata/geneseq/geneseq/AA1988.DAT.\*  
10: /SID52/gcgdata/geneseq/geneseq/AA1989.DAT.\*  
11: /SID52/gcgdata/geneseq/geneseq/AA1990.DAT.\*  
12: /SID52/gcgdata/geneseq/geneseq/AA1991.DAT.\*  
13: /SID52/gcgdata/geneseq/geneseq/AA1992.DAT.\*  
14: /SID52/gcgdata/geneseq/geneseq/AA1993.DAT.\*  
15: /SID52/gcgdata/geneseq/geneseq/AA1994.DAT.\*  
16: /SID52/gcgdata/geneseq/geneseq/AA1995.DAT.\*  
17: /SID52/gcgdata/geneseq/geneseq/AA1996.DAT.\*  
18: /SID52/gcgdata/geneseq/geneseq/AA1997.DAT.\*  
19: /SID52/gcgdata/geneseq/geneseq/AA1998.DAT.\*  
20: /SID52/gcgdata/geneseq/geneseq/AA1999.DAT.\*  
21: /SID52/gcgdata/geneseq/geneseq/AA2000.DAT.\*  
22: /SID52/gcgdata/geneseq/geneseq/AA2001.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1118	100.0	213	22	AA039624 Human polypeptide
2	1118	100.0	214	22	AA041410 Human polypeptide
3	826	73.9	175	21	AA034291 Human secreted pro
4	813	72.7	171	21	AA034292 Human secreted pro
5	731.5	65.4	200	21	AA033025 Zea mays protein f
6	725	64.8	216	21	AA014738 Arabidopsis thalia
7	725	64.8	216	21	AA018237 Arabidopsis thalia
8	725	64.8	218	21	AA014737 Arabidopsis thalia
9	725	64.8	218	21	AA018236 Arabidopsis thalia
10	705	63.1	158	21	AA033026 Zea mays protein f
11	663	59.3	143	21	AA033027 Zea mays protein f

12	661	59.1	143	21	AA014739 Arabidopsis thalia
13	661	59.1	143	21	AA018238 Arabidopsis thalia
14	652	58.3	143	21	AA028057 Arabidopsis thalia
15	602	53.8	132	21	AA028058 Arabidopsis thalia
16	338.5	30.3	132	21	AA025987 Zea mays protein f
17	88.5	7.9	510	19	AA069358 Pyridine nucleotid
18	82	7.3	729	21	AA044857 Mouse taste cell s
19	79.5	7.1	319	15	AA063296 Polypeptide encode
20	79.5	7.1	319	15	AA063297 Polypeptide encode
21	78	7.0	892	22	AA091553 C glutamicum prote
22	77.5	6.9	361	22	AA081007 Hydrogenase small
23	77.5	6.9	369	21	AA010120 Arabidopsis thalia
24	77.5	6.9	398	21	AA021667 Arabidopsis thalia
25	77.5	6.9	452	21	AA021666 Arabidopsis thalia
26	77.5	6.9	457	21	AA010119 Arabidopsis thalia
27	77.5	6.9	469	21	AA021665 Arabidopsis thalia
28	77.5	6.9	475	21	AA010118 Arabidopsis thalia
29	77	6.9	257	22	AA039519 Human polypeptide
30	76.5	6.8	753	22	AA094374 Human protein sequ
31	76	6.8	369	21	AA022172 Arabidopsis thalia
32	76	6.8	374	21	AA022171 Arabidopsis thalia
33	76	6.8	412	21	AA022170 Arabidopsis thalia
34	76	6.8	566	22	AA04762 Human vesicle traf
35	76	6.8	635	22	AA039243 Human polypeptide
36	76	6.8	660	22	AA041029 Human polypeptide
37	75.5	6.8	347	19	AA072010 HSV-2 strain SB5 C
38	75.5	6.8	548	19	AA072165 HSV-2 strain SB5 C
39	75.5	6.8	999	21	AA053685 Arabidopsis thalia
40	75.5	6.8	1008	21	AA053684 Arabidopsis thalia
41	75.5	6.8	1065	21	AA053683 Arabidopsis thalia
42	75.5	6.8	3011	16	AA067588 Hepatitis C virus
43	74.5	6.7	164	22	AA040762 Human polypeptide
44	74.5	6.7	500	21	AA084626 Amino acid sequenc
45	74.5	6.7	500	21	AA092051 HrpCa9 polypeptide

## ALIGNMENTS

RESULT 1  
AA039624  
ID AA039624 standard; Protein: 213 AA.  
XX  
AC AA039624;  
XX  
DT 22-OCT-2001 (first entry)  
XX  
DE Human polypeptide SEQ ID NO 2769.  
XX  
KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;  
KW peripheral nervous system; neuropathy; central nervous system; CNS;  
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;  
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;  
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;  
KW leukaemia.  
XX  
OS Homo sapiens.  
XX  
PN WO200153312-A1.  
XX  
PD 26-JUL-2001.  
XX  
PF 26-DEC-2000; 2000WO-US34263.  
XX  
PR 21-JAN-2000; 2000US-0488725.  
PR 25-APR-2000; 2000US-0552317.  
PR 09-JUL-2000; 2000US-0598042.  
PR 19-JUL-2000; 2000US-0620312.  
PR 03-AUG-2000; 2000US-0653450.  
PR 14-SEP-2000; 2000US-0662191.  
PR 19-OCT-2000; 2000US-0693036.  
PR 29-NOV-2000; 2000US-0727344.  
XX

(HYSE-) HYSEQ INC.

PA XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;  
 PI XX Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;  
 PI XX Zhao QA, Zhou P, Goodrich R, Drmanac RT;  
 XX XX WPI; 2001-442253/47.  
 DR DR N-PSDB; AAI58780.  
 XX XX Novel nucleic acids and polypeptides, useful for treating disorders  
 PT such as central nervous system injuries -  
 XX XX Example 4; SEQ ID NO 2769; 10078pp; English.  
 XX XX The invention relates to human nucleic acids (AAI57798-AAI61369) and  
 CC the encoded polypeptides (AAM3642-AAM42213) with nootropic,  
 CC immunosuppressant and cytostatic activity. The polynucleotides are useful  
 CC in gene therapy. A composition containing a polypeptide or polynucleotide  
 CC of the invention may be used to treat diseases of the peripheral nervous  
 CC system, such as peripheral nervous injuries, peripheral neuropathy and  
 CC localised neuropathies and central nervous system diseases, such as  
 CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic  
 CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the  
 CC utilisation of the activities such as: Immune system suppression,  
 CC Actin/inhibin activity, chemotactic/chemokinetic activity, haemostatic  
 CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,  
 CC assays for receptor activity, arthritis and inflammation, leukaemias and  
 CC C.N.S disorders.  
 CC Note: The sequence data for this patent did not form part of the printed  
 CC specification.  
 XX XX Sequence 213 AA;  
 XX XX

Query Match 100.0%; Score 1118; DB 22; Length 213;  
 Best Local Similarity 100.0%; Pred. No. 5.1e-117;  
 Matches 213; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAVLSAPGLRGFRILGRSSVGPVAVQARGVHQSVDTPGSSSTQPALPKARAVAPKPSRSRG 60  
 Db 1 mavlsapglrgfrilgrssvpgvargvhqsvatdpgsstqpalkparavapkpssrg 60  
 QY 61 EYVVAKLDDLVNWARSSSLWPMFTFGLACCAVEMHMAAPRYDMDRFGVVFASPRQSDVM 120  
 Db 61 eyvvaklddlvnwarssslwpmftfglaccavemhmmaaprydmdrfgvvfrsprqsdvm 120  
 QY 121 IVAGTLTNKMAPALRKVYDQMPPEPRVYVSMGSCANGGGYHYYSVVRGCDRIYVDIYI 180  
 Db 121 ivagtltnkmapalrkvydqmpeprvyvsmgscangggyyhysvvrsgcdriyvpdviyi 180  
 QY 181 PGCPTAEALLYGILQKIKRRLQIYWR 213  
 Db 181 pgcptaeallygilqlkikrrrlqiwyrr 213

RESULT 2  
 AAM41410  
 ID AAM41410 standard; Protein; 214 AA.  
 XX AC AAM41410;  
 XX XX 22-OCT-2001 (first entry)  
 DT XX Human polypeptide SEQ ID NO 6341.  
 DE XX Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;  
 XX KW peripheral nervous system; neuropathy; central nervous system; CNS;  
 KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;  
 KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;  
 KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;  
 KW leukaemia.  
 XX KW Homo sapiens.  
 OS

XX WO200153312-A1.  
 PN 26-JUL-2001.  
 PD 26-DEC-2000; 2000WO-US34263.  
 PF 21-JAN-2000; 2000US-0488725.  
 PR 25-APR-2000; 2000US-0552317.  
 PR 09-JUL-2000; 2000US-0598042.  
 PR 19-JUL-2000; 2000US-0620312.  
 PR 03-AUG-2000; 2000US-0653450.  
 PR 14-SEP-2000; 2000US-0662191.  
 PR 19-OCT-2000; 2000US-0693036.  
 PR 29-NOV-2000; 2000US-0727344.  
 XX (HYSE-) HYSEQ INC.  
 PA Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;  
 PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;  
 PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;  
 XX WPI; 2001-442253/47.  
 DR N-PSDB; AAI60566.  
 XX Novel nucleic acids and polypeptides, useful for treating disorders  
 PT such as central nervous system injuries -  
 XX XX Example 2; SEQ ID NO 6341; 10078pp; English.  
 XX XX The invention relates to human nucleic acids (AAI57798-AAI61369) and  
 CC the encoded polypeptides (AAM3642-AAM42213) with nootropic,  
 CC immunosuppressant and cytostatic activity. The polynucleotides are useful  
 CC in gene therapy. A composition containing a polypeptide or polynucleotide  
 CC of the invention may be used to treat diseases of the peripheral nervous  
 CC system, such as peripheral nervous injuries, peripheral neuropathy and  
 CC localised neuropathies and central nervous system diseases, such as  
 CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic  
 CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the  
 CC utilisation of the activities such as: Immune system suppression,  
 CC Actin/inhibin activity, chemotactic/chemokinetic activity, haemostatic  
 CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,  
 CC assays for receptor activity, arthritis and inflammation, leukaemias and  
 CC C.N.S disorders.  
 CC Note: The sequence data for this patent did not form part of the printed  
 CC specification.  
 XX XX Sequence 214 AA;  
 XX XX

Query Match 100.0%; Score 1118; DB 22; Length 214;  
 Best Local Similarity 100.0%; Pred. No. 5.2e-117;  
 Matches 213; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAVLSAPGLRGFRILGRSSVGPVAVQARGVHQSVDTPGSSSTQPALPKARAVAPKPSRSRG 60  
 Db 2 mavlsapglrgfrilgrssvpgvargvhqsvatdpgsstqpalkparavapkpssrg 61  
 QY 61 EYVVAKLDDLVNWARSSSLWPMFTFGLACCAVEMHMAAPRYDMDRFGVVFASPRQSDVM 120  
 Db 62 eyvvaklddlvnwarssslwpmftfglaccavemhmmaaprydmdrfgvvfrsprqsdvm 121  
 QY 121 IVAGTLTNKMAPALRKVYDQMPPEPRVYVSMGSCANGGGYHYYSVVRGCDRIYVDIYI 180  
 Db 122 ivagtltnkmapalrkvydqmpeprvyvsmgscangggyyhysvvrsgcdriyvpdviyi 181  
 QY 181 PGCPTAEALLYGILQKIKRRLQIYWR 213  
 Db 182 pgcptaeallygilqlkikrrrlqiwyrr 214

RESULT 3  
 AAB34291

ID XX AAB34291 standard; Protein; 175 AA.  
AC XX AAB34291;  
XX DT 02-FEB-2001 (first entry)  
DE XX Human secreted protein BLAST search protein SEQ ID NO: 137.  
DE XX  
DE XX  
KW Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;  
KW antiallergic; hepatotropic; antidiabetic; antiinflammatory; antitumor;  
KW vulnerary; anticonvulsant; antibacterial; antifungal; antiparasitic;  
KW cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;  
KW neurological disease; infection; human; secreted protein.  
XX OS Homo sapiens.  
XX PN WO200005352-A2.  
XX PD 21-SEP-2000.  
XX PF 09-MAR-2000; 2000WO-US06044.  
XX PR 12-MAR-1999; 99US-0124099.  
XX PR 03-DEC-1999; 99US-0168661.  
XX PA (HUMA-) HUMAN GENOME SCI INC.  
XX PI Rosen CA, Ruben SM, Komatsoulis G;  
XX DR WPI; 2000-602124/57.  
XX PT Novel human secreted proteins useful for diagnosis, prevention and  
PT treatment of disorders including neurological, cell proliferative,  
PT cardiovascular, autoimmune and inflammatory disorders and microbial  
PT infections  
XX PS Disclosure; Pages 379-380; 383pp; English.  
XX CC The invention relates to the isolation of genes AAC59507-C59556 encoding  
CC the human secreted proteins AAB34218-B34264. This sequence represents a  
CC peptide fragment homologous to the protein encoded by the gene isolated  
CC in the present invention. The sequence is a search result from a BLASTX  
CC homology search. The genes and proteins are useful for preventing,  
CC ameliorating or treating medical conditions, e.g. by protein or gene  
CC therapy. The genes are isolated from a range of human tissues disclosed  
CC in the specification. The nucleic acids, proteins, antibodies and  
CC (ant)agonists are useful in the diagnosis, treatment and prevention of:  
CC (a) cancer, e.g. breast and ovarian cancer, and other cancers of the  
CC adrenal gland, bone, bone marrow, breast, gastrointestinal tract, liver,  
CC lung, or urogenital; (b) immune disorders e.g. Addison's disease,  
CC allergies, autoimmune haemolytic anaemia, autoimmune thyroiditis,  
CC diabetes mellitus, Crohn's disease, multiple sclerosis, rheumatoid  
CC arthritis and ulcerative colitis; (c) cardiovascular disorders such  
CC as myocardial ischaemias; (d) wound healing; (e) neurological diseases  
CC as viral, bacterial, fungal and parasitic infections.  
XX SQ Sequence 175 AA;  
Query Match 73.9%; Score 826; DB 21; Length 175;  
Best Local Similarity 88.6%; Pred. No. 2e-84; Indels 4; Gaps 1;  
Matches 155; Conservative 7; Mismatches 9;  
QY 14 ILGLRSSVGPVQARGVHQSVATDGPSTQPALPKARAVAPK-----PSSRGEYVAKLDD 69  
Db 1 ilavrsygaalqvrghssmaadpsstqavsgatavvypkaalpsr:gyvvaakldd 60  
QY 70 LYNWARRSLWPTGLACCAVENHMAAPRYDMDRGVYFRASPRSDYMIAGILTNN 129  
Db 61 llnwarrslwptlglaccavennhmaaprydmdrgvfrasprsdymivagiltkn 120  
QY 130 MAPALRKVDQMPPEYRVVYVSMGSCANGGGYHYHSYVVRGCDRIVPVDIYIPGCP 184

Db 121 mapalrkvydgmpeyrvvsmgscanggggyhsyvsrvrgcdriypvdyvpgcp 175  
RESULT 4  
AAB34292  
ID AAB34292 standard; Protein; 171 AA.  
XX AC AAB34292;  
XX DT 02-FEB-2001 (first entry)  
XX DE Human secreted protein BLAST search protein SEQ ID NO: 138.  
XX DE  
XX KW Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;  
KW antiallergic; hepatotropic; antidiabetic; antiinflammatory; antitumor;  
KW vulnerary; anticonvulsant; antibacterial; antifungal; antiparasitic;  
KW cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;  
KW neurological disease; infection; human; secreted protein.  
XX OS Homo sapiens.  
XX PN WO200005352-A2.  
XX PD 21-SEP-2000.  
XX PF 09-MAR-2000; 2000WO-US06044.  
XX PR 12-MAR-1999; 99US-0124099.  
XX PR 03-DEC-1999; 99US-0168661.  
XX PA (HUMA-) HUMAN GENOME SCI INC.  
XX PI Rosen CA, Ruben SM, Komatsoulis G;  
XX DR WPI; 2000-602124/57.  
XX PT Novel human secreted proteins useful for diagnosis, prevention and  
PT treatment of disorders including neurological, cell proliferative,  
PT cardiovascular, autoimmune and inflammatory disorders and microbial  
PT infections  
XX PS Disclosure; Page 382; 383pp; English.  
XX CC The invention relates to the isolation of genes AAC59507-C59556 encoding  
CC the human secreted proteins AAB34218-B34264. This sequence represents a  
CC fragment of the protein encoded by the gene isolated in the present  
CC invention. The sequence is used as a query sequence for doing BLASTX  
CC searches to determine homologous sequence to the protein. The genes and  
CC proteins are useful for preventing, ameliorating or treating medical  
CC conditions, e.g. by protein or gene therapy. The genes are isolated  
CC from a range of human tissues disclosed in the specification. The  
CC nucleic acids, proteins, antibodies and (ant)agonists are useful in the  
CC diagnosis, treatment and prevention of: (a) cancer, e.g. breast and  
CC ovarian cancer, and other cancers of the adrenal gland, bone, bone  
CC marrow, breast, gastrointestinal tract, liver, lung, or urogenital;  
CC (b) immune disorders e.g. Addison's disease, allergies, autoimmune  
CC haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus,  
CC Crohn's disease, multiple sclerosis, rheumatoid arthritis and  
CC ulcerative colitis; (c) cardiovascular disorders such as myocardial  
CC ischaemias; (d) wound healing; (e) neurological diseases e.g. cerebral  
CC anoxia and epilepsy; and (f) infectious diseases such as viral,  
CC bacterial, fungal and parasitic infections.  
XX SQ Sequence 171 AA;  
Query Match 72.7%; Score 813; DB 21; Length 171;  
Best Local Similarity 91.8%; Pred. No. 5.5e-83;  
Matches 157; Conservative 0; Mismatches 14; Indels 0; Gaps 0;  
QY 14 ILGLRSSVGPVQARGVHQSVATDGPSTQPALPKARAVAPKPSRGEYVAKLDDLVNW 73  
|||||

Db 1 ilgirssvqavqargvhqsvardgsstqpalpkaravapkpssrgeyyvavklldlvnw 60  
QY 74 ARRSLSWMTFGLACCAVENWMAAPRYDMDFGVVFRASPRQSDVMIVAGTLTKMAPA 133  
Db 61 arrrsllwmtfglaccavemhnaaprydmndrfgvvfraspqsdvmivagtltnkmapa 120  
QY 134 LRKYVDQMPPEPRYVWSMGSCANGGYYHYHSYSVVRCGDRIVPDVYIPGCP 184  
Db 121 lrkyvdqmppepryvvsmgscanxxxxxxxcccccccccdripvdyipgcp 171

RESULT 5  
AAG33025  
ID AAG33025 standard; Protein; 200 AA.  
XX AC AAG33025;  
XX DT 17-OCT-2000 (first entry)  
XX ZEa mays protein fragment SEQ ID NO: 39954.  
KW Protein identification; signal transduction pathway; metabolic pathway;  
KW hybridisation assay; genetic mapping; gene expression control; promoter;  
KW termination sequence; corn.  
XX OS Zea mays subsp. mays.  
XX PN EP1033405-A2.  
XX PD 06-SEP-2000.  
XX PF 25-FEB-2000; 2000EP-0301439.  
XX PR 05-FEB-1999; 99US-0121825.  
PR 05-MAR-1999; 99US-0123180.  
PR 09-MAR-1999; 99US-0123548.  
PR 23-MAR-1999; 99US-0125788.  
PR 25-MAR-1999; 99US-0126264.  
PR 29-MAR-1999; 99US-0126785.  
PR 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128234.  
PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0129845.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 23-APR-1999; 99US-0130891.  
PR 28-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 04-MAY-1999; 99US-0132407.  
PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 07-MAY-1999; 99US-0132487.  
PR 11-MAY-1999; 99US-0132863.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 02-JUL-1999; 99US-0142154.  
PR 06-JUL-1999; 99US-0142055.  
PR 08-JUL-1999; 99US-0142390.  
PR 09-JUL-1999; 99US-0142803.  
PR 12-JUL-1999; 99US-0142920.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145089.  
PR 23-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.

```
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160788.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161982.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 65.4%; Score 731.5; DB 21; Length 200;
Best Local Similarity 73.2%; Pred No. 9.1e-74;
Matches 134; Conservative 17; Mismatches 31; Indels 1; Gaps 1;

QY 31 HQSVATDGPSTQPALPKARAVAPKPSRGEYVYVAKLDLVLNNARRSSLWPMTFTGLACCA 90

99US-0147935.
99US-0148171.
99US-0148319.
99US-0148341.
99US-0148565.
99US-0148684.
99US-0149368.
99US-0149175.
99US-0149426.
99US-0149722.
99US-0149723.
99US-0149929.
99US-0149902.
99US-0149930.
99US-0150566.
99US-0150884.
99US-0151065.
99US-0151066.
99US-0151080.
99US-0151303.
99US-0151438.
99US-0151930.
99US-0152363.
99US-0153070.
99US-0153758.
99US-0154018.
99US-0154039.
99US-0154779.
99US-0155139.
99US-0155486.
99US-0155659.
99US-0156458.
99US-0156596.
99US-0157117.
99US-0157753.
99US-0157865.
99US-0158029.
99US-0158232.
99US-0158369.
99US-0159293.
99US-0159294.
99US-0159295.
99US-0159329.
99US-0159330.
99US-0159331.
99US-0159637.
99US-0159638.
99US-0159584.
99US-0160741.
99US-0160767.
99US-0160788.
99US-0160770.
99US-0160814.
99US-0160815.
99US-0160980.
99US-0160981.
99US-0160989.
99US-0161404.
99US-0161405.
99US-0161406.
99US-0161359.
99US-0161360.
99US-0161361.
99US-0161920.
99US-0161982.
99US-0161993.
99US-0162142.

Db 19 ysaaaaasptspa-pyggappamstaeafvsvskvddlmwarkswipmtftglacca 77
QY 91 VEMHMAAPRYDMDFGVYFRASPRQSDVMIVAGTLTNKMAPALRKVYDQMPEPRVYVSM 150
Db 78 vemhagasrydfdrfgvifprsprqsdcmivaglttnkmapalrkvydqnpqprwvism 137
QY 151 GSCANGGGYHYHSYSVVRGCDRIYIPVDIYIPGCPPTAEALLYGILQORKIKRERLQIW 210
Db 138 gscanggyhyhsyvsrvrgcdriypvdiypvgopptaeallyglqkknrrkdfhw 197
QY 211 YRR 213
Db 198 wtk 200

RESULT 6
AAG14738
ID AAG14738 standard; Protein; 216 AA.
XX
AC AAG14738;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 14714.
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126284.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
```

PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139452.  
PR 17-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139452.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 24-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 18-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0147038.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.

PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0148968.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 15-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 14-OCT-1999; 99US-0159638.  
PR 14-OCT-1999; 99US-0159584.  
PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160768.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 21-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160980.  
PR 22-OCT-1999; 99US-0160981.  
PR 22-OCT-1999; 99US-0160989.  
PR 25-OCT-1999; 99US-0161404.  
PR 25-OCT-1999; 99US-0161405.  
PR 25-OCT-1999; 99US-0161406.  
PR 26-OCT-1999; 99US-0161359.  
PR 26-OCT-1999; 99US-0161360.  
PR 26-OCT-1999; 99US-0161361.  
PR 28-OCT-1999; 99US-0161920.  
PR 28-OCT-1999; 99US-0161992.  
PR 28-OCT-1999; 99US-0161993.  
PR 29-OCT-1999; 99US-0162142.



Query Match	64.8%; Score 725; DB 21; Length 216;
Best Local Similarity	68.2%; Pred. No. 5.4e-73;
Matches 135; Conservative 20; Mismatches 35; Indels 8; Gaps 2;	
QY 24	AVCARGV---HOSVATDGPSS-----TOPALPKARAVAPKPSRRGEYVVAKLDELVNKAR 75
DB 19	avaaavshlhtslpalspststptstyrtpgpbstppppglskaaeviskvddimwar 78
QY 76	RSLWPMTFTGLACCAVEMMHMAAPRYMDRGVFRASPROSDYVMIVAGTLTNKMAPALR 135
DB 79	tgsiwpmtftglaccavemmhmtgaarydlrfgilfrpsrgsdmivagtltnkmapalr 138
QY 136	KVYDQMPERYVYVSMGSCANGGYXHYYSVVRGCDRIVPVDIYIPGCPPFAEALLYGIL 195
DB 139	kyvdqmperyvismgscanggyynysvvrvgcdrivpvdiyyvgcptaecallygll 198
QY 196	QLQRIKRRRLQIWMYR 213
DB 199	qlqkkinrrkdfilhwink 216
RESULT 7	
AAI818237	
ID	AAI818237 standard; Protein; 216 AA.
XX	
AC	AAI818237;
XX	
DT	17-OCT-2000 (first entry)
XX	
DE	Arabidopsis thaliana protein fragment SEQ ID NO: 19564.
XX	
KW	Protein identification; signal transduction pathway; metabolic pathway;
KW	hybridisation assay; genetic mapping; gene expression control; promoter;
KW	termination sequence.
XX	
OS	Arabidopsis thaliana.
PN	
XX	
PD	EP1033405-A2.
XX	
PF	06-SEP-2000.
XX	
PF	25-FEB-2000; 2000EP-0301439.
XX	
PF	99US-0121825.
PR	25-FEB-1999;
PR	99US-0123180.
PR	05-MAR-1999;
PR	99US-0123548.
PR	09-MAR-1999;
PR	99US-0125788.
PR	23-MAR-1999;
PR	99US-0126264.
PR	25-MAR-1999;
PR	99US-0125785.
PR	01-APR-1999;
PR	99US-0127462.
PR	06-APR-1999;
PR	99US-0128234.
PR	08-APR-1999;
PR	99US-0128714.
PR	16-APR-1999;
PR	99US-0123845.
PR	19-APR-1999;
PR	99US-0130077.
PR	21-APR-1999;
PR	99US-0130449.
PR	23-APR-1999;
PR	99US-0130510.
PR	23-APR-1999;
PR	99US-0130891.
PR	28-APR-1999;
PR	99US-0131449.
PR	30-APR-1999;
PR	99US-0132048.
PR	30-APR-1999;
PR	99US-0132407.
PR	04-MAY-1999;
PR	99US-0132484.
PR	05-MAY-1999;
PR	99US-0132485.
PR	06-MAY-1999;
PR	99US-0132486.
PR	06-MAY-1999;
PR	99US-0132487.
PR	07-MAY-1999;
PR	99US-0132863.
PR	11-MAY-1999;
PR	99US-0134256.
PR	14-MAY-1999;
PR	99US-0134218.
PR	14-MAY-1999;
PR	99US-0154219.
PR	14-MAY-1999;
PR	99US-0134221.
PR	14-MAY-1999;
PR	99US-0134370.
PR	18-MAY-1999;
PR	99US-0134768.
PR	19-MAY-1999;
PR	99US-0134941.
PR	20-MAY-1999;
PR	99US-0135124.
PR	21-MAY-1999;
PR	99US-0135353.
PR	21-MAY-1999;
PR	99US-0135621.
PR	24-MAY-1999;
PR	99US-0136021.
PR	25-MAY-1999;
PR	99US-0136392.
PR	27-MAY-1999;
PR	99US-0136782.
PR	28-MAY-1999;
PR	99US-0137222.
PR	01-JUN-1999;
PR	99US-0137528.
PR	03-JUN-1999;
PR	99US-0137502.
PR	04-JUN-1999;
PR	99US-0137724.
PR	07-JUN-1999;
PR	99US-0138094.
PR	08-JUN-1999;
PR	99US-0138540.
PR	10-JUN-1999;
PR	99US-0138847.
PR	10-JUN-1999;
PR	99US-0139119.
PR	14-JUN-1999;
PR	99US-0139452.
PR	16-JUN-1999;
PR	99US-0139453.
PR	16-JUN-1999;
PR	99US-0139452.
PR	17-JUN-1999;
PR	99US-0139454.
PR	18-JUN-1999;
PR	99US-0139455.
PR	18-JUN-1999;
PR	99US-0139456.
PR	18-JUN-1999;
PR	99US-0139457.
PR	18-JUN-1999;
PR	99US-0139458.
PR	18-JUN-1999;
PR	99US-0139459.
PR	18-JUN-1999;
PR	99US-0139460.
PR	18-JUN-1999;
PR	99US-0139461.
PR	18-JUN-1999;
PR	99US-0139462.
PR	18-JUN-1999;
PR	99US-0139463.
PR	18-JUN-1999;
PR	



PR 14-MAY-1999; 99US-0134321.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 17-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139753.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143342.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144532.  
PR 20-JUL-1999; 99US-0144864.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.

PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 05-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 06-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149829.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 16-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 14-OCT-1999; 99US-0159638.  
PR 18-OCT-1999; 99US-0159584.  
PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160768.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 21-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160980.  
PR 22-OCT-1999; 99US-0160981.  
PR 22-OCT-1999; 99US-0160989.  
PR 25-OCT-1999; 99US-0161404.

4

PR	22-JUL-1999;	99US-01451192.
PR	23-JUL-1999;	99US-01451145.
PR	23-JUL-1999;	99US-01452118.
PR	26-JUL-1999;	99US-01452224.
PR	26-JUL-1999;	99US-01452276.
PR	27-JUL-1999;	99US-0145913.
PR	27-JUL-1999;	99US-0145918.
PR	27-JUL-1999;	99US-0145919.
PR	28-JUL-1999;	99US-0145951.
PR	02-AUG-1999;	99US-0146386.
PR	02-AUG-1999;	99US-0146388.
PR	02-AUG-1999;	99US-0146389.
PR	03-AUG-1999;	99US-0147038.
PR	04-AUG-1999;	99US-0147204.
PR	04-AUG-1999;	99US-0147302.
PR	05-AUG-1999;	99US-0147132.
PR	05-AUG-1999;	99US-0147260.
PR	06-AUG-1999;	99US-0147303.
PR	06-AUG-1999;	99US-0147416.
PR	09-AUG-1999;	99US-0147433.
PR	09-AUG-1999;	99US-0147935.
PR	10-AUG-1999;	99US-0148171.
PR	11-AUG-1999;	99US-0148319.
PR	12-AUG-1999;	99US-0148341.
PR	13-AUG-1999;	99US-0148565.
PR	13-AUG-1999;	99US-0148684.
PR	16-AUG-1999;	99US-0149368.
PR	17-AUG-1999;	99US-0149175.
PR	18-AUG-1999;	99US-0149426.
PR	20-AUG-1999;	99US-0149722.
PR	20-AUG-1999;	99US-0149723.
PR	20-AUG-1999;	99US-0149929.
PR	23-AUG-1999;	99US-0149902.
PR	23-AUG-1999;	99US-0149930.
PR	25-AUG-1999;	99US-0150566.
PR	26-AUG-1999;	99US-0150884.
PR	27-AUG-1999;	99US-0151065.
PR	27-AUG-1999;	99US-0151066.
PR	30-AUG-1999;	99US-0151080.
PR	30-AUG-1999;	99US-0151303.
PR	31-AUG-1999;	99US-0151438.
PR	01-SEP-1999;	99US-0151930.
PR	07-SEP-1999;	99US-0152363.
PR	10-SEP-1999;	99US-0153070.
PR	13-SEP-1999;	99US-0153758.
PR	15-SEP-1999;	99US-0154018.
PR	16-SEP-1999;	99US-0154039.
PR	20-SEP-1999;	99US-0154779.
PR	22-SEP-1999;	99US-0155139.
PR	23-SEP-1999;	99US-0155486.
PR	24-SEP-1999;	99US-0155659.
PR	28-SEP-1999;	99US-0156458.
PR	29-SEP-1999;	99US-0156596.
PR	04-OCT-1999;	99US-0157117.
PR	05-OCT-1999;	99US-0157753.
PR	06-OCT-1999;	99US-0157865.
PR	07-OCT-1999;	99US-0158029.
PR	08-OCT-1999;	99US-0158232.
PR	12-OCT-1999;	99US-0158369.
PR	13-OCT-1999;	99US-0159293.
PR	13-OCT-1999;	99US-0159294.
PR	14-OCT-1999;	99US-0159329.
PR	14-OCT-1999;	99US-0159330.
PR	14-OCT-1999;	99US-0159331.
PR	14-OCT-1999;	99US-0159637.
PR	18-OCT-1999;	99US-0159638.
PR	21-OCT-1999;	99US-0159584.
PR	21-OCT-1999;	99US-0160741.
PR	21-OCT-1999;	99US-0160767.
PR	21-OCT-1999;	99US-0160770.
PR	21-OCT-1999;	99US-0160814.
PR	22-JUL-1999;	99US-01451192.
PR	23-JUL-1999;	99US-01451145.
PR	23-JUL-1999;	99US-01452118.
PR	26-JUL-1999;	99US-01452224.
PR	26-JUL-1999;	99US-01452276.
PR	27-JUL-1999;	99US-0145913.
PR	27-JUL-1999;	99US-0145918.
PR	27-JUL-1999;	99US-0145919.
PR	28-JUL-1999;	99US-0145951.
PR	02-AUG-1999;	99US-0146386.
PR	02-AUG-1999;	99US-0146388.
PR	02-AUG-1999;	99US-0146389.
PR	03-AUG-1999;	99US-0147038.
PR	04-AUG-1999;	99US-0147204.
PR	04-AUG-1999;	99US-0147302.
PR	05-AUG-1999;	99US-0147132.
PR	05-AUG-1999;	99US-0147260.
PR	06-AUG-1999;	99US-0147303.
PR	06-AUG-1999;	99US-0147416.
PR	09-AUG-1999;	99US-0147433.
PR	09-AUG-1999;	99US-0147935.
PR	10-AUG-1999;	99US-0148171.
PR	11-AUG-1999;	99US-0148319.
PR	12-AUG-1999;	99US-0148341.
PR	13-AUG-1999;	99US-0148565.
PR	13-AUG-1999;	99US-0148684.
PR	16-AUG-1999;	99US-0149368.
PR	17-AUG-1999;	99US-0149175.
PR	18-AUG-1999;	99US-0149426.
PR	20-AUG-1999;	99US-0149722.
PR	20-AUG-1999;	99US-0149723.
PR	20-AUG-1999;	99US-0149929.
PR	23-AUG-1999;	99US-0149902.
PR	23-AUG-1999;	99US-0149930.
PR	25-AUG-1999;	99US-0150566.
PR	26-AUG-1999;	99US-0150884.
PR	27-AUG-1999;	99US-0151065.
PR	27-AUG-1999;	99US-0151066.
PR	30-AUG-1999;	99US-0151080.
PR	30-AUG-1999;	99US-0151303.
PR	31-AUG-1999;	99US-0151438.
PR	01-SEP-1999;	99US-0151930.
PR	07-SEP-1999;	99US-0152363.
PR	10-SEP-1999;	99US-0153070.
PR	13-SEP-1999;	99US-0153758.
PR	15-SEP-1999;	99US-0154018.
PR	16-SEP-1999;	99US-0154039.
PR	20-SEP-1999;	99US-0154779.
PR	22-SEP-1999;	99US-0155139.
PR	23-SEP-1999;	99US-0155486.
PR	24-SEP-1999;	99US-0155659.
PR	28-SEP-1999;	99US-0156458.
PR	29-SEP-1999;	99US-0156596.
PR	04-OCT-1999;	99US-0157117.
PR	05-OCT-1999;	99US-0157753.
PR	06-OCT-1999;	99US-0157865.
PR	07-OCT-1999;	99US-0158029.
PR	08-OCT-1999;	99US-0158232.
PR	12-OCT-1999;	99US-0158369.
PR	13-OCT-1999;	99US-0159293.
PR	13-OCT-1999;	99US-0159294.
PR	14-OCT-1999;	99US-0159329.
PR	14-OCT-1999;	99US-0159330.
PR	14-OCT-1999;	99US-0159331.
PR	14-OCT-1999;	99US-0159637.
PR	18-OCT-1999;	99US-0159638.
PR	21-OCT-1999;	99US-0159584.
PR	21-OCT-1999;	99US-0160741.
PR	21-OCT-1999;	99US-0160767.
PR	21-OCT-1999;	99US-0160770.
PR	21-OCT-1999;	99US-0160814.
PR	22-JUL-1999;	99US-01451192.
PR	23-JUL-1999;	99US-01451145.
PR	23-JUL-1999;	99US-01452118.
PR	26-JUL-1999;	99US-01452224.
PR	26-JUL-1999;	99US-01452276.
PR	27-JUL-1999;	99US-0145913.
PR	27-JUL-1999;	99US-0145918.
PR	27-JUL-1999;	99US-0145919.
PR	28-JUL-1999;	99US-0145951.
PR	02-AUG-1999;	99US-0146386.
PR	02-AUG-1999;	99US-0146388.
PR	02-AUG-1999;	99US-0146389.
PR	03-AUG-1999;	99US-0147038.
PR	04-AUG-1999;	99US-0147204.
PR	04-AUG-1999;	99US-0147302.
PR	05-AUG-1999;	99US-0147132.
PR	05-AUG-1999;	99US-0147260.
PR	06-AUG-1999;	99US-0147303.
PR	06-AUG-1999;	99US-0147416.
PR	09-AUG-1999;	99US-0147433.
PR	09-AUG-1999;	99US-0147935.
PR	10-AUG-1999;	99US-0148171.
PR	11-AUG-1999;	99US-0148319.
PR	12-AUG-1999;	99US-0148341.
PR	13-AUG-1999;	99US-0148565.
PR	13-AUG-1999;	99US-0148684.
PR	16-AUG-1999;	99US-0149368.
PR	17-AUG-1999;	99US-0149175.
PR	18-AUG-1999;	99US-0149426.
PR	20-AUG-1999;	99US-0149722.
PR	20-AUG-1999;	99US-0149723.
PR	20-AUG-1999;	99US-0149929.
PR	23-AUG-1999;	99US-0149902.
PR	23-AUG-1999;	99US-0149930.
PR	25-AUG-1999;	99US-0150566.
PR	26-AUG-1999;	99US-0150884.
PR	27-AUG-1999;	99US-0151065.
PR	27-AUG-1999;	99US-0151066.
PR	30-AUG-1999;	99US-0151080.
PR	30-AUG-1999;	99US-0151303.
PR	31-AUG-1999;	99US-0151438.
PR	01-SEP-1999;	99US-0151930.
PR	07-SEP-1999;	99US-0152363.
PR	10-SEP-1999;	99US-0153070.
PR	13-SEP-1999;	99US-0153758.
PR	15-SEP-1999;	99US-0154018.
PR	16-SEP-1999;	99US-0154039.
PR	20-SEP-1999;	99US-0154779.
PR	22-SEP-1999;	99US-0155139.
PR	23-SEP-1999;	99US-0155486.
PR	24-SEP-1999;	99US-0155659.
PR	28-SEP-1999;	99US-0156458.
PR	29-SEP-1999;	99US-0156596.
PR	04-OCT-1999;	99US-0157117.
PR	05-OCT-1999;	99US-0157753.
PR	06-OCT-1999;	99US-0157865.
PR	07-OCT-1999;	99US-0158029.
PR	08-OCT-1999;	99US-0158232.
PR	12-OCT-1999;	99US-0158369.
PR	13-OCT-1999;	99US-0159293.
PR	13-OCT-1999;	99US-0159294.
PR	14-OCT-1999;	99US-0159329.
PR	14-OCT-1999;	99US-0159330.
PR	14-OCT-1999;	99US-0159331.
PR	14-OCT-1999;	99US-0159637.
PR	18-OCT-1999;	99US-0159638.
PR	21-OCT-1999;	99US-0159584.
PR	21-OCT-1999;	99US-0160741.
PR	21-OCT-1999;	99US-0160767.
PR	21-OCT-1999;	99US-0160770.
PR	21-OCT-1999;	99US-0160814.

Query Match

Best Local Similarity

Mismatches

Conservative

Score

DB

Length

Indels

Gaps

2;

QY

24

AVQARGV---HQSVDGDPSS-----TOPALPKARAVAPKPSRGGEYVAKLDOLYNWAR

75

Db

21

avaaasvshlhtslpalspsstsytrpgpptspppglskaaeviskvdlnnwar

80

QY

76

RSSLWPMTEGLACCAVEMHMAAPRYDMDFGVVFRASPRQSDVMIVAGLTNKNMAPALR

135

Db

81

tgslwpmctfglaccavemhmtgaarydlrfgliifrspqscdmivagltlnkmapalr

140

QY

136

KYVDQMPPEYVVMGSCANGGYHYHSYVVRGCDRIVPDYIYPCGPTAEALLYGIL

195

Db

141

kvydgmpeprwvismgscanggyhyhsyvsrgcdripvdiypvpcptaeallygl

200

QY

196

QLQRKIKRERRLQIWYRR

213

Db

201

qliqkinrrkdlhwnk

218

RESULT

10

AAG33026

standard; Protein: 158 AA.

XX

AAG33026;

XX

17-OCT-2000 (first entry)

DE

Zea mays protein fragment SEQ ID NO: 39955.

KW

Protein identification; signal transduction pathway; metabolic pathway;

KW

hybridisation assay; genetic mapping; gene expression control; promoter;

XX

termination sequence; corn.

OS

Zea mays subsp. mays.

XX

EP1033405-A2.

PD

06-SEP-2000.

XX

25-FEB-2000; 2000EP-0301439.

XX

25-FEB-1999; 99US-0121825.

PR

05-MAR-1999; 99US-0123160.

PR

09-MAR-1999; 99US-0123548.

PR

23-MAR-1999; 99US-0125788.

PR

25-MAR-1999; 99US-0126264.

PR

25-MAR-1999; 99US-0126785.

PR

01-APR-1999; 99US-0127462.

PR

06-APR-1999; 99US-0128234.

PR

08-APR-1999; 99US-0128714.

PR

16-APR-1999; 99US-0129845.

PR

19-APR-1999; 99US-0130077.

PR

21-APR-1999; 99US-0130449.

PR

23-APR-1999; 99US-0130510.

PR

23-APR-1999; 99US-0130891.

PR

28-APR-1999; 99US-0131449.

PR 30-APR-1999; 99US-0132048.  
PR 30-APR-1999; 99US-0132407.  
PR 04-MAY-1999; 99US-0132484.  
PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 06-MAY-1999; 99US-0132487.  
PR 07-MAY-1999; 99US-0132863.  
PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 18-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 18-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144684.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 14-OCT-1999; 99US-0159638.  
PR 18-OCT-1999; 99US-0159584.

```
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 21-OCT-1999; 99US-0160880.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 22-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 25-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161922.
PR 28-OCT-1999; 99US-0161933.
PR 29-OCT-1999; 99US-0162142.
PR 29-OCT-1999; 99US-0162143.

Query Match 63.18; Score 705; DB 21; Length 158;
Best Local Similarity 79.68; Pred. No. 6.1e-71;
Matches 125; Conservative 15; Mismatches 17; Indels 0; Gaps 0;

QY 57 SSRGEYVYAKLDLNVARRSSLPMTFGLACCAVEMMHMAAPRYDMDFGVVFRASPRQ 116
DQ 2 sktaefvsvkvdldmnmwarksgipmtfglaccavemmhagasydfdrigvfrsprq 61

QY 117 SDVMTVAGTLNKNAPALRYVDQMPFRYVYVMSGANGGGYHYHSYSVYRGCDRIVPV 176
DQ 62 scdmivagtlknmapalrkvydqmpeprwvismgscanggggyhyhsysvvrqcdriwpv 121

QY 177 DIYIPGCPPTAALLYGLQRLQRIKRLQIWR 213
DQ 122 diypgcpptaeallyglvqlqklnrkdflhwtk 158

RESULT 11
ID AAG33027 standard; Protein; 143 AA.
XX AAG33027;
AC AAG33027;
DT 17-OCT-2000 (first entry)
DE Zea mays protein fragment SEQ ID NO: 39956.
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence; corn.
OS Zea mays subsp. mays.
PN EP1033405-A2.
PD 06-SEP-2000.
XX 25-FEB-2000; 2000EP-0301439.
XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127452.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.

PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140891.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143824.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
```

```
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 25-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146388.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 08-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158022.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159294.
PR 14-OCT-1999; 99US-0159322.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.

PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 26-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 28-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 59.3%; Score 663; DB 21; Length 143;
Best Local Similarity 81.8%; Pred. No. 2.7e-66;
Matches 117; Conservative 12; Mismatches 14; Indels 0; Gaps 0;

QY 71 VNWARSSLPMTFTGLACCAVEMHMAAPRYDMDFGVVFRASPRQSDVMIIVAGTLTNKM 130
Db 1 mnwarkgslwpmftglaccavemhmagasrydfgrvfrpsrqscmivagtltnkm 60

QY 131 APALRKVDQPEPRVVSVMGSCANGGGYHYHSYVVRGCDRIVPVDIYIPGCPPTAEAL 190
Db 61 apalrkvydqmpeprvismgscangggvyhsyvsvrgcdrivpvdiypvgcpptaeal 120

QY 191 LYGILQLOKIKRRERLQIWR 213
Db 121 lygvlqlqkkinrrkdfnlhwtk 143

RESULT 12
AAG14739
ID AAG14739 standard; Protein; 143 AA.
XX
AC AAG14739;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 14715.
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence.
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126284.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
```



PR 23-APR-1999; 99US-0130891.  
PR 28-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 30-APR-1999; 99US-0132407.  
PR 04-MAY-1999; 99US-0132484.  
PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 06-MAY-1999; 99US-0132487.  
PR 07-MAY-1999; 99US-0132863.  
PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136031.  
PR 27-MAY-1999; 99US-0136382.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140384.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142134.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 09-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0130884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 02-AUG-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159233.  
PR 13-OCT-1999; 99US-0159284.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.

```
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161320.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.
PR 29-OCT-1999; 99US-0162142.

Query Match 59.1%; Score 661; DB 21; Length 143;
Best Local Similarity 81.1%; Pred. No. 4.5e-66;
Matches 116; Conservative 13; Mismatches 14; Indels 0; Gaps 0;

QY 71 VNWARRSLWPMTFGLACCAVENMHAAPRYDMDREGVFRASPROSDVMIVAGTLTNKM 130
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 1 mwwartsipmtfglaccavemmhgaaryldirgfrspqscdmvagtctkm 60

QY 131 APALRKVDMPPEPRYVYVMSGANGGYHYHSYSVVRGCDRVPVDIYPOCPPTAEAL 190
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 61 apalrkvydmppeprvysngscanggyhyhsysvvrsgcdripvdiyvpcpptaeeal 120

QY 191 LYGILQLOKRIKRRRLQIWR 213
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 121 lygilqlqkxrkrkrlqihwvk 143

RESULT 13
AAG18238
ID AAG18238 standard; Protein; 143 AA.
XX AC AAG18238;
XX DT 17-OCT-2000 (first entry)
XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 19565.
XX KW Protein identification; signal transduction pathway; metabolic pathway;
XX KW hybridisation assay; genetic mapping; gene expression control; promoter;
XX KW termination sequence.
XX OS Arabidopsis thaliana.
XX PN EP1033405-A2.
XX PD 06-SEP-2000.
XX PF 25-FEB-2000; 2000EP-0301439.
XX PR 25-FEB-1999; 99US-0121825.
XX PR 05-MAR-1999; 99US-0123180.
XX PR 09-MAR-1999; 99US-0123548.
XX PR 23-MAR-1999; 99US-0125788.
XX PR 25-MAR-1999; 99US-0126264.
XX PR 29-MAR-1999; 99US-0126785.
XX PR 01-APR-1999; 99US-0127462.
XX PR 06-APR-1999; 99US-0128234.
XX PR 08-APR-1999; 99US-0128714.
XX PR 16-APR-1999; 99US-0129845.
XX PR 19-APR-1999; 99US-0130077.
XX PR 21-APR-1999; 99US-0130449.
```

```
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
```

PR 20-JUL-1999; 99US-0144532.  
PR 20-JUL-1999; 99US-0144584.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 22-JUL-1999; 99US-0145194.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145313.  
PR 27-JUL-1999; 99US-0145318.  
PR 27-JUL-1999; 99US-0145319.  
PR 28-JUL-1999; 99US-0145351.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149802.  
PR 23-AUG-1999; 99US-0149930.  
PR 23-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0156559.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0158293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.

PR 14-OCT-1999; 99US-0159637.  
PR 14-OCT-1999; 99US-0159638.  
PR 18-OCT-1999; 99US-0159584.  
PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160768.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 21-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160980.  
PR 22-OCT-1999; 99US-0160981.  
PR 22-OCT-1999; 99US-0160989.  
PR 23-OCT-1999; 99US-0161404.  
PR 23-OCT-1999; 99US-0161405.  
PR 25-OCT-1999; 99US-0161406.  
PR 26-OCT-1999; 99US-0161359.  
PR 26-OCT-1999; 99US-0161360.  
PR 26-OCT-1999; 99US-0161361.  
PR 28-OCT-1999; 99US-0161920.  
PR 28-OCT-1999; 99US-0161992.  
PR 28-OCT-1999; 99US-0161993.  
PR 28-OCT-1999; 99US-0162142.  
  
Query Match 59.1%; Score 661; DB 21; Length 143;  
Best Local Similarity 81.1%; Pred. No. 4.5e-66;  
Matches 116; Conservative 13; Mismatches 14; Indels 0; Gaps 0;  
  
QY 71 VNARRSLWPTFGLACCAVENMMHMAAPRYDMREGVVRASPRQSDVMIVAGTLTNKM 130  
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
Db 1 mnwargtsiwpmtfglaccavemhtgaarydlidrfgiifrsprqsdcmivagtltnkm 60  
  
QY 131 APALRKVYDQMPPEPRVYVSMGSCANGGYYHYYSVVRGCDRIVPDVYIPGCPPTAEAL 190  
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
Db 61 apalrkvydqmpeprvysmgscangggyyhyysvvrsgcdriypvdiypgcpptaeal 120  
  
QY 191 LYGILQLRKIKRRRLQIWR 213  
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
Db 121 lygilqlqkkinrrkdfhlwnk 143  
  
RESULT 14  
AAG28057  
ID AAG28057 standard; Protein; 143 AA.  
XX AC AAG28057;  
XX DT 17-OCT-2000 (first entry)  
XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 33132.  
XX KW Protein identification; signal transduction pathway; metabolic pathway;  
KW hybridisation assay; genetic mapping; gene expression control; promoter;  
KW termination sequence.  
XX OS Arabidopsis thaliana.  
XX PN EP1033405-A2.  
XX PD 06-SEP-2000.  
XX PF 25-FEB-2000; 2000EP-0301439.  
XX PR 25-FEB-1999; 99US-0121825.  
PR 05-MAR-1999; 99US-0123180.  
PR 09-MAR-1999; 99US-0123548.  
PR 23-MAR-1999; 99US-0125788.  
PR 25-MAR-1999; 99US-0126264.  
PR 29-MAR-1999; 99US-0126785.  
PR 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128234.  
PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0129845.  
PR 19-APR-1999; 99US-0130077.

PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 23-APR-1999; 99US-0130691.  
PR 28-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 30-APR-1999; 99US-0132407.  
PR 04-MAY-1999; 99US-0132484.  
PR 05-MAY-1999; 99US-0132485.  
PR 08-MAY-1999; 99US-0132486.  
PR 08-MAY-1999; 99US-0132487.  
PR 07-MAY-1999; 99US-0132863.  
PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134270.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0136094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143621.  
PR 15-JUL-1999; 99US-0144008.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 12-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156598.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159330.